Structure and Bonding 185
Series Editor: D. Michael P. Mingos

D. Michael P. Mingos
Paul R. Raithby *Editors*

21st Century Challenges in Chemical Crystallography I

History and Technical Developments

MOREMEDIA
Springer

185
Structure and Bonding

Aims and Scope

Structure and Bonding is a publication which uniquely bridges the journal and book format. Organized into topical volumes, the series publishes in depth and critical reviews on all topics concerning structure and bonding. With over 50 years of history, the series has developed from covering theoretical methods for simple molecules to more complex systems.

Topics addressed in the series now include the design and engineering of molecular solids such as molecular machines, surfaces, two dimensional materials, metal clusters and supramolecular species based either on complementary hydrogen bonding networks or metal coordination centers in metal-organic framework materials (MOFs). Also of interest is the study of reaction coordinates of organometallic transformations and catalytic processes, and the electronic properties of metal ions involved in important biochemical enzymatic reactions.

Volumes on physical and spectroscopic techniques used to provide insights into structural and bonding problems, as well as experimental studies associated with the development of bonding models, reactivity pathways and rates of chemical processes are also relevant for the series.

Structure and Bonding is able to contribute to the challenges of communicating the enormous amount of data now produced in contemporary research by producing volumes which summarize important developments in selected areas of current interest and provide the conceptual framework necessary to use and interpret mega-databases.

We welcome proposals for volumes in the series within the scope mentioned above. Structure and Bonding offers our authors and readers:

- OnlineFirst publication. Each chapter is published online as it is finished, ahead of the print volume
- Wide dissemination. The chapters and the volume will be available on our platform SpringerLink, one of the largest collections of scholarly content in the world. SpringerLink attracts more than 50 million users at 15.000 institutions worldwide.
- Easy manuscript preparation. Authors do not have to spend their valuable time on the layout of their contribution. Springer will take care of all the layout related issues and will provide support throughout the complete process.

More information about this series at http://www.springer.com/series/430

D. Michael P. Mingos • Paul R. Raithby
Editors

21st Century Challenges in Chemical Crystallography I

History and Technical Developments

With contributions by

D. R. Allan · C. M. Beavers · S. J. Coles · R. I. Cooper · L. E. Hatcher · S. J. W. Holgate · D. M. P. Mingos · S. A. Moggach · I. D. H. Oswald · A. R. Pallipurath · P. R. Raithby · L. K. Saunders · J. M. Skelton · S. J. Teat · C. A. Tovee · M. R. Warren

Springer

Editors
D. Michael P. Mingos
Inorganic Chemistry Laboratory
University of Oxford
Oxford, UK

Paul R. Raithby
Department of Chemistry
University of Bath
Bath, UK

ISSN 0081-5993 ISSN 1616-8550 (electronic)
Structure and Bonding
ISBN 978-3-030-64742-1 ISBN 978-3-030-64743-8 (eBook)
https://doi.org/10.1007/978-3-030-64743-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The Centenary of von Laue and Bragg's discovery and interpretation of the diffraction of X-ray by crystals prompted several Festschrifts celebrating the wonderful achievements of X-ray crystallography during the last century. To a large extent, the solving of the structures of small molecules has become relatively straightforward and is accessible to many chemists who do not have long or mature knowledge of the fundamentals of the subject. Although the danger of errors being made by less experienced practitioners has not been completely eliminated, the software used to process the raw data and refine the structures has become more sophisticated and now provides additional steps to limit the introduction of obvious mistakes in space group assignments and the introduction of doubtful procedures. These changes also raise the important question regarding the directions in which small molecule crystallography should be taking in the coming century. Paul Raithby and I decided it would be useful to have a volume of *Structure and Bonding* summarising recent developments in small molecule X-ray crystallography, identify problems that may be encountered when using the technique and explore new techniques that point towards future directions for the subject. These include the studies of structures as they evolve under light irradiation, with high pressures or under a range of different environmental conditions. We were also fortunate that several world leading experts in the subject have agreed to provide chapters for this volume.

This introductory chapter on *The Early History of X-Ray Crystallography* by Michael Mingos recounts the development of X-ray crystallography at the beginning of the twentieth century by von Laue and the Braggs. The technique they discovered did not enable scientists to look at the molecular world by looking through a more powerful microscope, but it provided data that when processed enabled scientists to calculate the structures of molecules and appreciate their three-dimensional structures. It provided the zeitgeist of our time, and the belief that knowledge of the structure leads to a more profound understanding of the function and properties of that class of molecule has influenced the developments of all sciences from physics to medicine. It is a fascinating technique because unlike the optical microscope it required the development of a deeper understanding of the way in which the X-rays

interact with the electron density in the planes of the crystal and the development of models in order to model this electron density satisfactorily. This chapter traces how these problems were overcome. In the early days, the structures of even simple organic molecules would take a Ph.D. student several months or even years to solve the structure. In time and particularly since the 1950s, the development of more sophisticated equipment and the massive rise in computing power made it possible to solve the three-dimensional structure of an organic molecule within a few minutes with the latest detectors on a laboratory instrument. This successful trajectory has resulted in the ability to study evermore complex molecules and use smaller and smaller crystals. The structures of over a million organic and organometallic compounds are now archived in the most commonly used database and this wealth of information creates a new set of problems for future generations of scientists.

Recent Developments in Refinement and Analysis of X-Ray Crystal Structures has been assembled by Richard Cooper. This chapter discusses the development and testing of structural models which accurately represent disordered crystal structures and guest molecules. The interpretation and control of structural geometry and displacement parameters are also discussed. For routine structure determinations, the independent atom model (IAM) is sufficient to explain X-ray scattering with enough accuracy that atom positions and displacements can be confidently determined. This structure analysis paradigm has not changed significantly since the early days of crystallography, nor have the underlying mathematical procedures. Nevertheless, increased computing power and optimised algorithms for linear algebra and Fourier transforms have increased the practical limits of the complexity that can be handled within a reasonable timescale.

The chapter *Leading Edge Chemical Crystallography Service Provision and Its Impact on Crystallographic Data Science in the Twenty-First Century* is discussed by Simon Coles, David Allan, Christine Beavers, Simon Teat, Stephen Holgate and Clare Tovee. The authors are responsible for national crystallographic services and cyclotron facilities in the UK and the USA. National facilities provide state-of-the-art crystallographic instrumentation and processes and tend to act as an indicator for the direction of travel for the community. This chapter discusses the future development of national facilities including those using synchrotron radiation sources and also addresses the challenges of harnessing the development of the resulting large databases so that they are able to drive new science in areas such as crystal engineering. The chapter provides insight into how specific aspects of crystallography are currently developing and shows how they can increasingly interact or integrate with other areas. This increased inter-operation will provide a much richer methodology and enable crystallography to be a key component in a broad range of research long into the future. The main message of this review is that chemical crystallography has the potential to do much more. Taking a more data-integrated, or even data-centric, approach it can be a leader in chemical and materials science research in the longer term. The discipline does, nevertheless, need to embrace different mindsets and give appropriate training for working in the changing

laboratory conditions, adopting new approaches to "publishing" and introducing new data science ways of undertaking research.

The chapter by Iain Oswald and Stephen Moggach *Reviews the Crystallographic Analysis of Crystals Under High Pressure Conditions.* Recent advances in the field are discussed with a focus on both organic and inorganic systems. It is an area of crystallography that has seen a rapid expansion over the last two decades. Advances in technology and data processing have facilitated the discovery of new materials, polymorphs and chemistries under extreme conditions. The occurrence and characterisation of polymorphs are of particular concern for the pharmaceutical industry. They discuss these advances using examples of organic and metal–organic materials as well as providing guidance to the pitfalls to be avoided while conducting these studies.

Jonathan Skeleton, Lauren Hatcher, Mark Warren, Anuradha Pallipurath and Lucy Saunders' have contributed the chapter *Watching Photochemistry Happen: Recent Developments in Dynamic Single-Crystal X-Ray Diffraction Studies*. Mechanistic information on solid-state photochemical reactions has traditionally come from spectroscopy and qualitative or quantitative modelling. The crystallographic contribution was limited to snapshots of endpoints and long-lived intermediates. Recent advances in X-ray sources and detectors have made it possible to follow solid-state reactions in situ with dynamic single-crystal X-ray diffraction (SCXRD) methods, allowing a full set of atomic positions to be determined over the course of the reaction. These experiments provide valuable structural information that can be used to interpret spectroscopic measurements and to inform materials design and optimisation. Paul Raithby has contributed a chapter on *Time Resolved Single-Crystal X-Ray Crystallography* which traces the development of this technique from its beginnings more than 30 years ago. The importance of being able to "watch" chemical processes as they occur rather than just be limited to three-dimensional pictures of the reactant and final product is emphasised, and time resolved crystallography provides the opportunity to bring the dimension of time into the crystallographic experiment. The technique has evolved in time with developments in technology: synchrotron radiation, cryoscopic techniques, tuneable lasers, increased computing power and vastly improved X-ray detectors. The shorter the lifetime of the species being studied, the more complex is the experiment. The chapter focusses on the results of solid-state reactions that are activated by light since this process does not require the addition of a reagent to the crystalline material and the single-crystalline nature of the solid may be preserved.

The sequel volume *Twenty-first Century Challenges in Chemical Crystallography II: Structural Correlations and Data Interpretation* covers the way in which the structural information obtained from chemical crystallography has been utilised and interpreted. The chapter headings are given below:

1. Mingos: *Historical Development of Historical Correlations*
2. Grabowsky et al: *The Advent of Quantum Crystallography: Form and Structure Factors from Quantum Mechanics for Advanced Refinement and Wavefunction Fitting*

3. Overgaad et al: *Experimental Charge Densities from Multipole modelling: Moving into the twenty-first Century*
4. Macgregor et al: *Computational Studies of the Solid-State Molecular Organometallic (SMOM) Chemistry of Rhodium Alkane Complexes*

Oxford, UK
Bath, UK
August 2020

D. M. P. Mingos
P. R. Raithby

Contents

Struct Bond (2020) 185: 1–42
https://doi.org/10.1007/430_2020_73

Published online: 4 October 2020

Early History of X-Ray Crystallography

D. Michael P. Mingos

Contents

Abstract The discovery of the optical microscope played an important role in the scientific revolution of the seventeenth century because it enabled one to directly view objects which were invisible to the naked eye. In 1667 Robert Hooke improved the microscope invented in the previous century in Holland and used it to examine the "microscopic" appearance of snowflakes and plants. Others were able to view for themselves the presence of very small objects and the structures of plants, hair, skin, bones etc. The development of X-ray crystallography at the beginning of the twentieth century by von Laue and the Braggs played an equally important role in the scientific revolution which has shaped our lives. The technique they discovered did not enable scientists to look at the molecular world by looking through a more powerful microscope, but it provided data which when processed enabled scientists to calculate the structures of molecules and appreciate their three-dimensional structures. It provided the zeitgeist of our time that the knowledge of the structure

D. M. P. Mingos (✉)
Inorganic Chemistry Laboratory, Oxford University, Oxford, UK
e-mail: Michael.mingos@seh.ox.ac.uk

would lead to a more profound understanding of the function and properties of that class of molecule.

This chapter recounts the early history of the development of this important technique and describes how the early technical problems were overcome. It is a fascinating technique because unlike the optical microscope it required the development of a deeper understanding of the way in which the X-rays interact with the electron density in the planes of the crystal and the development of models in order to model this electron density satisfactorily. This chapter traces how these problems were overcome. In the early days, the structures of even simple organic molecules would take a PhD student several months or even years to solve the structure. In time and particularly since the 1950s, the development of more sophisticated equipment and the massive rise in computing power made it possible to solve the three-dimensional structure of an organic molecule within a few minutes with the latest detectors on a laboratory instrument. This successful trajectory has resulted in the ability to study ever more complex molecules and use smaller and smaller crystals. The structures of over a million organic and organometallic compounds are now archived in the most commonly used database, and this wealth of information creates a new set of problems for future generations of scientists.

Keywords Bragg equation · Crystals · Databases · Diffraction · Fourier series

1 Introduction

Structure and Bonding has been published by Springer for more than 50 years, and as its title suggests, it seeks to publish reviews associated with the structure of chemical compounds and their interpretation using current models of chemical bonding [1]. However, we have never published a volume devoted completely to the most important structural technique of modern chemistry, i.e. X-ray crystallography. Paul Raithby and I decided it would be timely to publish a volume of *Structure and Bonding*, which not only celebrates the great advances made in this technique over the last 110 years but highlights issues which remain problematic. It also seeks to identify new emerging areas which will enable the subject to continue growing at the same rapid rate that it has achieved over the last century. This growth has been possible by effective interdisciplinary collaborations between physicists, chemists, biologists, mathematicians, computer programmers and engineers who have worked together to enlarge the technical capabilities of the discipline and provide the equations and programmes to enable the conversion of the raw data into an accurate description of the molecule. They developed new X-ray sources with more monochromatic and focussed beams and improved the measurements of the diffraction spots, sped up the acquisition of data and the conversion of the data into accurate structural information and improve the knowledge of the fundamental

physics underlying the technique. Most importantly the subject has attracted successive generations of talented scientists who wanted to answer the really big scientific problems of the age and were willing to push the subject to new limits and develop techniques to study ever more complex systems. It was inconceivable that when the structure of sodium chloride and zinc sulphide was initially solved in 1912 that within 50 years the same technique would be used to solve the structures of myoglobin and haemoglobin and provide information concerning the structure of the DNA which would revolutionise the field of genetics.

This extraordinary successful scientific endeavour has to a large extent made the solving of small molecules structures routine, but the solution of many structures within hours carries with it the responsibility to ensure that the structural information is accurate and reproduceable and archived in a fashion which makes it accessible to the whole scientific community. As the technology has developed, opportunities have arisen to undertake crystallographic experiments that were not possible a few years previously. These include the studies of structures as they are irradiated by light, exposed to high pressures or other environmental conditions. The question of crystallographic disorder has remained a continuing issue with structural determinations and cannot always be handled satisfactorily using automatic structure solution programmes. Distinguishing atoms which are adjacent in the periodic table and very light atoms such as hydrogen have always been a problem for the technique since X-ray scattering and diffraction depend on the number of electrons associated with the atom. The use of quantum mechanical calculations to model the electron densities in atoms and molecules has developed greatly in recent decades and enabled scientists the opportunity to use this technique to investigate bonding issues in molecules and solids [2, 3]. This introductory chapter introduces the historical background to the discovery and exploitation of this technique. The subsequent chapters written by world experts in small molecule crystallography discuss recent developments and the prospects for the future. The extension of the technique to molecules of biological systems has proved to be particularly successfully and has resulted in many important and significant insights and will not be covered in this volume, and the reader is directed to other excellent sources [4]. Examples of excellent textbooks which present a more detailed account of the theoretical principles of X-ray diffraction and its practical aspects are given in references [5–14].

2 Early Experiments

In 1895 Wilhelm Conrad Röntgen [15] became the first person to detect and appreciate the novel properties of X-rays, and he started the sequence of experiments which led to the development of X-ray crystallography. Several centuries earlier the discovery of the optical microscope enabled scientists to directly view objects which were invisible to the naked eye and contributed greatly to the scientific revolution. In 1667 Robert Hooke improved the microscope first developed in 1590 by two Dutch spectacle makers, Hans and Zacharias Janssen, and he explored and illustrated the

structures of snowflakes and plants and introduced the public to the widespread occurrence of fleas and lice and their detailed appearance. Hooke's study of the bark of the cork tree revealed its microscopic structure and resulted in the discovery of the *cell*, which proved to be the building block of all life forms. He also used his microscope to show that ancient cells were present in fossilised wood. He concluded that fossils had once been living creatures whose cells had become mineralised. He also concluded that some species that had once existed must have become extinct. He thereby started a process which was to lead to Darwin's theory of evolution in 1865. In 1665 at the age of 30, Hooke published *Micrographia*, the first ever scientific bestseller, which illustrated his understanding of nature and light, his highly developed skills in designing and constructing scientific instruments and his skills as an artist. Antonie van Leeuwenhoek, who used a microscope with one lens to observe insects and other specimens, was the first to observe bacteria. Despite many innovative developments, it did not prove possible in the subsequent two centuries to use the increased power of microscopes to examine the structures at the atomic level of metals, nonmetals and the compounds they formed when they combined. The discovery of X-rays at the end of the nineteenth century did not immediately enable scientists to look directly at the structures of molecules. Nevertheless, the diffraction patterns produced on photographic plates when crystals were exposed to X-rays enabled scientists to mathematically manipulate the data and calculate the three-dimensional arrangements of atoms in crystals for the first time. This procedure now described as X-ray crystallography has had an enormous impact on physics, chemistry, biology and medicine during the last century [5–14].

Röntgen's discovery occurred accidentally in his laboratory in Wurzburg, Germany, where he was trying to establish whether cathode rays could pass through glass [15]. Cathode rays observed in discharge tubes were first discovered in 1867 and were shown to be streams of electrons by J.J. Thomson in 1897 [16] and led to his determination of the *e*/*m* ratio of the electron particle. In the darkened room, Röntgen observed a glow coming from a nearby chemically coated screen. He attributed this glow to a novel but not previously studied form of radiation, which he described as X-rays, and he set about to define their true nature by further scientific studies. These established that X-rays unlike cathode rays are electromagnetic energy waves that act similar to light rays but have wavelengths approximately 1,000 times shorter than those of visible light. The absence of a charge and their higher energies meant that they could penetrate human flesh and muscles but not higher-density substances such as bone. The significance of this observation was rapidly recognised by the medical profession and has proved to be the most important application of his discovery for humanity ever since. The X-rays were unable to penetrate heavier metals such as lead, and this property was used to develop effective protection procedures for doctors and physicists who subsequently developed the medical and scientific applications.

In 1912 Max von Laue, at the Institute of Theoretical Physics in Munich, became aware of his colleague's Paul Ewald's theoretical research on the optical properties of a solid containing a regular arrangement of resonators and argued that a crystalline solid consists of a three-dimensional arrangement of atoms and molecules whose

interatomic distances may correspond approximately to the wavelengths of X-rays and may generate a three-dimensional diffraction pattern when radiated with X-rays. He persuaded Walter Friedrich and Paul Knipping, who had previous experience with a more powerful X-ray bulb than available to Röntgen and were able to produce a collimated and narrower primary X-ray beam, to undertake some preliminary experiments [17, 18]. Using large single crystals of hydrated copper sulphate and zinc sulphide, they showed that when the beam of X-rays strikes the crystal the incident beam is scattered in many directions and the resultant diffraction pattern could be observed using a photographic plate (after it had been developed using standard photographic reagents). In this experiment the X-rays were not directly being used as microscope in the manner described earlier by Hooke, but the observation of a diffraction pattern suggested the possibility of obtaining information about the separations of a regular atomic grid within the crystal [17, 18]. This required a clearer understanding of the physics responsible for the diffraction process and the development of mathematical equations which could be used to convert the diffraction pattern into a model which may give information about the positions of the atoms. The major obstacle in the pathway leading from the observed diffraction pattern to the desired crystal structure is known as the phase problem and is discussed further below.

The interference patterns observed in these experiments supported the characterisation of X-rays as electromagnetic waves with very short wavelengths rather than as corpuscles. Von Laue with his able coworkers' demonstration of the phenomenon of X-ray diffraction proved to be inspirational. Remarkably within the next decade and even with the interruptions caused by the outbreak of World War I, these initial observations were developed into an important scientific tool which contributed to the second scientific revolution. This proved to be an important landmark in modern science because it enabled chemists, physicists and molecular biologists to obtain a unique understanding of the fundamental structures of molecules, minerals, metals, alloys, proteins and enzymes at atomic resolutions. It underpinned what was to become the *zeitgeist* of the age, i.e. establishing the structures of molecules leads to an understanding of their function. More germane to the present volume of *Structure and Bonding*, the structural information it provided was used to understand the principles of chemical bonding and led to the deeper understanding of physical and chemical properties of important key building blocks of physics, chemistry and biology. Such an important discovery has of course attracted many scholars, and it is not possible to summarise all the relevant contributions in a short review. References [19–27] give more detailed accounts and personal reminiscences. The important early papers on the diffraction of X-rays have also been collated in two volumes [25, 26]. My brief summary has drawn extensively on the excellent book by Andre Authier – *Early Days of X-ray Crystallography* published in 2013 [21].

Friedrich and Knipping experiments and Laue's initial interpretation published in June 1912 had an immediate impact within the scientific community [21]. Max Planck recalled that scientists in Berlin "felt that a remarkable feat had been achieved", and Albert Einstein defined the experiment as "among the most glorious that physics has seen so far". Not surprisingly this paper received a good reception

from traditional optical crystallographers since the well-defined spots and their regular disposition on the photographic plate were seen as conclusive evidence that atoms are arranged regularly in a space-lattice configuration in crystals and suggested the possibility of atomic resolution. As Alfred Tutton, the English crystallographer, overstated in November 1912, "the space-lattice structure of crystals . . . is now rendered visible to our eyes" [21]. Within a year of the discovery of X-ray diffraction, William Lawrence Bragg had reinterpreted the phenomenon responsible for the diffraction pattern and determined the first unit cell dimensions of simple salts, and within 2 years he and his father (William Henry Bragg) had reported the first crystal structure determinations [19, 20, 23, 24]. The discoveries of von Laue and Bragg gave birth to two new sciences, X-ray crystallography and X-ray spectroscopy, and two Nobel Prizes: Max von Laue "for his discovery of the diffraction of X-rays by crystals" in 1914 and to the Braggs "for their services in the analysis of crystal structure by X-ray Crystallography in 1915 (see Table 1).

Von Laue's experiment had confirmed unambiguously the wave nature of X-rays and suggested that the wavelength was sufficiently short that it gave rise to a diffraction pattern from the target crystal. However, he did not completely or succinctly articulate the physics which was responsible for the phenomenon. He argued that the internal order of atoms within the crystal, which had been hinted at for several centuries, results in a three-dimensional diffraction grating with spacings much smaller than that commonly associated with visible light and their exposure to X-rays resulted in the observed diffraction pattern. Specifically, the lack of knowledge concerning the nature of the lattice and the wavelength(s) of the X-rays limited his contribution. Nonetheless, this seed of an idea rapidly grew and blossomed, and it was able to draw on the circumstantial evidence which had been obtained previously from optical crystallography which had been developed initially to study mineralogical samples but was also used to study inorganic crystals made by chemists in the nineteenth century. This cross fertilisation is discussed in more detail in the subsequent section.

Table 1 gives an indication of the enormous subsequent impact of von Laue's contribution by listing the Nobel Laureates who were recognised for their contributions to the development of this important structural technique. In the early days, the winners were physicists who contributed to our understanding of the basic diffraction phenomenon. Specifically, the contributions of Arthur Compton and Louis de Broglie resolved the wave/particle duality of small atomic particles which resulted from the quantum mechanical description pioneered by Planck and Einstein and demonstrated that such particles were capable of giving diffraction patterns through their wave nature. Their insight was verified by Davisson and Thomson's demonstration of diffraction patterns obtained when an electron beam was passed through a crystal of nickel. Peter Debye defined the importance of thermal vibrations of atoms in crystals, and he showed that microcrystalline samples also produced diffraction patterns and thereby established X-ray powder diffraction, which proved to be so important in metallurgy, materials science and solid-state chemistry. Although Bragg's initial contributions were made studying crystalline samples, he encouraged his students to adopt a broader appreciation that the technique may be applied to

Table 1 Nobel Prizes directly associated with X-ray crystallography

Year	Laureates	Prize	Citation
1914	M. von Laue	Physics	"Diffraction of X-rays by crystals"
1915	W.H. and W.L. Bragg	Physics	"The analysis of crystal structures by means of X-rays"
1927	A.H. Compton	Physics	For showing that X-ray quanta scattered by electrons had longer wavelengths, the surplus energy having been transferred to the electrons – the Compton effect – demonstrated the particle properties of electromagnetic radiation
1929	L-V. de Broglie	Physics	Contribution to wave/particle duality and the equation which defines the wavelength of quantum particles such as electrons, neutrons and protons
1936	Petrus J. W. Debye	Physics	"For his contributions to our knowledge of molecular structure through his investigations on dipole moments and on the diffraction of X-rays and electrons in gases"
1937	C.J. Davisson and G.P. Thomson	Physics	"Or their experimental discovery of the diffraction of electrons by crystals"
1946	James B. Sumner	Chemistry	"For his discovery that enzymes can be crystallised"
1962	M.F. Perutz and J.C. Kendrew	Chemistry	"The structures of globular proteins"
1962	J.D. Watson, F.H.C. Crick and M. Wilkins	Medicine	"The molecular structure of nucleic acids and its significance for information transfer in living matter"
1964	D. Hodgkin	Chemistry	"X-ray techniques for the structures of important biochemical substances"
1972	S. Moore and W.H. Stein	Chemistry	"The chemical structures and catalytic activity of the active centre of the ribonuclease molecule"
1976	W.N. Lipscomb	Chemistry	"The structures of boranes illuminating problems of chemical bonding"
1982	A. Klug	Chemistry	"Development of crystallographic electron microscopy and the structural elucidation of nucleic acid-protein complexes used methods from X-ray diffraction, microscopy and structural modelling"
1985	J. Karle and H.A. Hauptman	Chemistry	"Developing direct methods for the determination of crystal structures"
1988	J. Deisenhofer, H. Michel and R. Huber	Chemistry	"Three-dimensional structure of a photosynthetic reaction centre"
1944	C. Shull and B. Brockhouse	Physics	Neutron diffraction and neutron inelastic scattering
1997	J. E. Walker	Chemistry	"Elucidation of the enzymatic mechanism underlying the synthesis of adenosine triphosphate (ATP)"
2003	R. MacKinnon and Peter Agre	Chemistry	"Structural and mechanistic studies of ion channels and water channels"

(continued)

Table 1 (continued)

Year	Laureates	Prize	Citation
2006	R. Kornberg	Chemistry	"Molecular basis of eukaryotic transcription"
2009	A.E. Yonath, T.A. Steitz and V. Ramakrishnan	Chemistry	"The structure and function of the ribosome"
2011	D. Shechtman	Chemistry	"Discovery of quasicrystals"
2012	B. Kobilka	Chemistry	"Structural studies of the G-protein-coupled receptors"
2017	J. Dubochet, J. Frank and R. Henderson	Chemistry	"For developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution"

For further biographical details of the Nobel Laureates and longer descriptions of their contribution go to http://nobelprize.org/nobel_prizes/chemistry/laureates/19xx/index.html inserting the appropriate year 19xx or 20xx or substitute chemistry for physics

biological systems. Specifically, John Bernal and William Astbury were encouraged to study the diffraction patterns of proteins. Sumner established in 1946 that enzymes produced diffraction patterns, and in subsequent years, the majority of Nobel Prizes summarised in Table 1 were awarded for the application of crystallographic techniques to biological systems. These contributions led to the establishment of molecular biology, and some individual contributions are summarised in Table 1. Despite their outstanding importance for the subsequent development of science, they will not be discussed in detail in this review, because the emphasis of this volume is on small molecule crystallography. Their inclusion in Table 1 does nonetheless emphasise the way in which technical and computing techniques advances have made it possible to extend a technique which initially was used for simple crystalline salts which have been developed to solve the structures of very large proteins and enzymes. The enormous amount of new structural data created by X-ray structural determinations also resulted in the award of many Nobel Prizes, which did not mention X-rays in the citation, but nonetheless would not have been possible without the structural information and insight obtained from this technique. These contributions will be discussed in a subsequent chapter of this volume. The enormous amount of new structural data which resulted from this technique and the related fields of X-ray powder diffraction, neutron diffraction and electron diffraction have created its own new problems, and the procedures which have been taken to deal with the archiving of the data and their subsequent accession are discussed in the chapter by Coles et al.

3 Optical Crystallography

The discoverers of X-ray diffraction were able to draw on the knowledge and expertise which had been developed by geologists and mineralogists over several centuries in universities, but the study of crystals has a much longer history. The occasional discovery of crystalline materials by primitive man was probably looked upon as a generous gift from the Gods, and they were incorporated into beautiful jewellery. Their colours, magnetic properties and sharp edges were utilised to enrich life, simplify daily tasks and enable them to explore the larger world. The properties of quartz in the natural state led the Greeks and Romans to define the more general term "crystal", and they were attracted by their obvious symmetry properties which resonated with the mathematically derived regularity of the Platonic solids. The mathematical basis of crystallography therefore had its origins in ancient Greece and Rome, and Pliny devoted five entire volumes of his work *Naturalis Historia* (77 CE) to the classification of "earths, metals, stones and gems". He described many minerals not known to the Greeks and discussed their applications and properties. He laid the fundamentals of crystallography by discussing crystal habits and was particularly taken by the octahedral shape of diamond crystals. It developed into a rigorous scientific discipline in the Renaissance with the discovery of the microscope which enabled mineralogists to examine and measure in a scientific manner a wide range of minerals and appreciate the great beauty of hexagonal symmetries in the diverse range of ice crystals. This eventually formed the basis of the following qualitative classification of crystalline minerals in the nineteenth century [28]:

1. Predominant crystal faces (prism-prismatic, pyramid-pyramidal and pinacoid-platy)
2. Crystal forms (cubic, octahedral, dodecahedral)
3. Aggregation of crystals or aggregates (fibrous, botryoidal, radiating, massive)
4. Crystal appearance (foliated/lamellar (layered), dendritic, bladed, acicular, lenticular, tabular (tablet shaped))

Perhaps the most influential text in the nineteenth century was the *Manual of Mineralogy* by James Dwight Dana, first published in 1848. The fourth edition was entitled *Manual of Mineralogy and Lithology* (1887), and it became the standard college text. It has been successively revised and updated in its subsequent 23 editions.

The atomistic view of the universe associated with Democritus (370 BCE) led to the proposal that crystals were made up of indivisibly small regularly shaped and identical "units". The Pythagorean school had defined around 550 BCE the five regular Platonic solids which bore some resemblance to the crystalline forms found in nature. The problem was related to the principles familiar to masons, who had successfully built pyramids in ancient Egypt and South America. Aristotle speculated that a crystal could be assembled if very small polyhedral "bricks" were packed together without leaving any space between, thereby avoiding vacuum cavities. In two dimensions all space could be tiled in this way using triangles, rectangles and

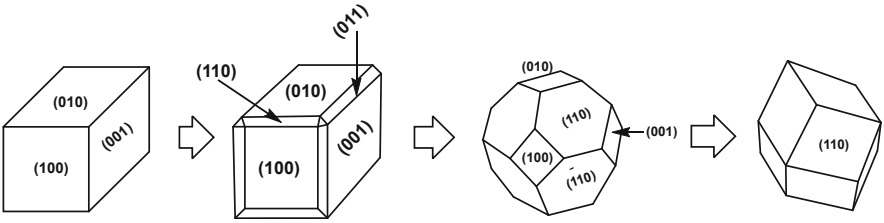


Fig. 1 An illustration of how crystal morphology may evolve as it grows. Starting as a cube it evolves into a cube with truncated edges into a rhombic dodecahedron with truncated corners and finally a rhombic dodecahedron

hexagons, and they speculated that in three dimensions construction blocks based on the Platonic solids could achieve the same result. Mathematically it was shown that pentagons could not completely tile 2D space and consequently Platonic solids with fivefold symmetry (icosahedron and the dodecahedron) were excluded as potential building blocks for crystals, leaving tetrahedra, cubes and octahedra as the remaining polyhedral building blocks. Kepler in the early seventeenth century rigorously established that the complete 3D space filling could be achieved by the cube, the octahedron, the rhomb-dodecahedron and the hexagonal prism. The observation that calcite could be fractured into smaller rhombohedra provided support for this view. Haüy argued that repeated cleavage would lead ultimately to the smallest and fundamental crystallite that could not be cleaved further without altering its essential physical and chemical nature, and this evolved into what we now describe as the unit cell [21].

Guglielmini and Carpeller at the end of the seventeenth century were the first to suggest that crystals shapes were related to chemical composition. For example, common salt was characterised by a cube, vitriol by a rhombohedral parallelepiped, nitre by a hexagonal prism and alum by an octahedron. Later studies demonstrated that the same chemical salt could crystallise in more than one modification and the significant and measurable feature of a crystal was not its shape but the "constancy of its interfacial angles". It also became apparent that the shape of a crystal could change as it grew (see Fig. 1). The accurate measurement of the interfacial angles was made possible by the construction of optical goniometers by Wollaston based on the crystal being mounted on a rotating spindle at the focal point of an adapted microscope. These measurements spawned the field of optical crystallography by Westfield, Gahn, Bergmann, Haüy and Pryce. These measurements also led to the realisation that crystals of different but related compounds could be isomorphic (Mitscherlich 1819), i.e. belong to identical crystal systems, but with slightly different dimensions [28–36]. These properties had important implications for the development of the periodic table because salts of compounds belonging to the same column of the periodic table not only had similar formulae but frequently proved to be isomorphic because their underlying atomic structures had similar symmetries although they have different cell dimensions because of changes in the sizes of the

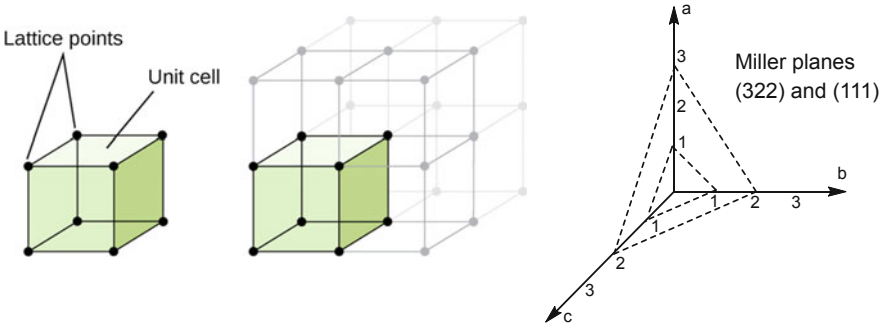


Fig. 2 A cubic cell, lattice and illustrations of the (111) and (322) Miller planes are shown on the right

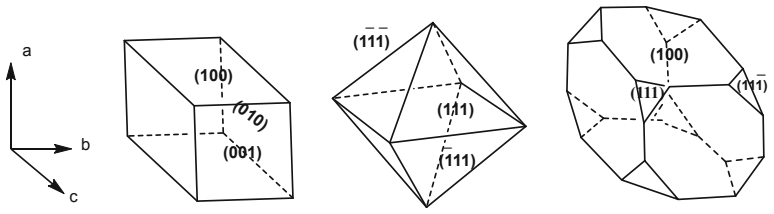


Fig. 3 Examples of the characterisation of crystal faces using Miller indices for a cube, octahedron and truncated cube

constituent atoms [37]. The interfacial angles could be related to the fundamental building block which is the unit cell shown in Fig. 2 [28–36].

The measurement of interfacial angles in crystals required a classification and notation for each face. This was achieved with a reference system based on three Cartesian axes collinear with the three axes of the unit cell. The crystal faces were associated with planes defined by multiple lengths of the unit cell (see Fig. 2). Weiss developed the mathematics underlying this methodology, and his student Neumann first suggested the use of the reciprocal of the lattice lengths to define the direction of the faces of the crystal planes by defining their intercepts with the axes of the crystal form. Neumann and Weiss's law of rational intercepts states that there exists a set of crystal axes which permits a specific crystal face to be characterised in terms of intercepts of the face with these axes. Each crystal face is thereby characterised by three indices (Miller indices). For those planes parallel to one of the crystal axes, its intercept is infinity, and hence its reciprocal is 0. If a face cuts a crystal axis along the negative axis, then this is indicated by a negative Miller index. Specific examples illustrating the Miller indices for a cubic and octahedral crystal are illustrated in Fig. 3. Figure 1 gives the Miller indices of the faces of a crystal as its morphology evolves from cubic to rhombohedral as discussed above. The results of the optical crystallographic measurements were generally represented by stereographic presentations, and the plane's reflection angle was assigned the appropriate Miller index.

Table 2 Crystal systems

Crystal system	Essential symmetry	Restrictions on lengths and angles of unit cell
Triclinic	None	None
Monoclinic	A twofold rotation axis and/or a mirror plane	$a \neq b \neq c$; $\alpha = \gamma = 90°$; $\beta \neq 90°$
Orthorhombic	Three twofold rotations and/or mirrors	$a \neq b \neq c$; $\alpha = \beta = \gamma = 90°$
Tetragonal	One fourfold rotation	$a = b \neq c$; $\alpha = \beta = \gamma = 90°$
Trigonal	One threefold rotation	$a = b = c$; $\alpha = \beta = 90°$; $\gamma = 120°$
Hexagonal	One sixfold rotation	$a = b \neq c$; $\alpha = \beta = 90°$; $\gamma = 120°$
Cubic	Four threefold rotation axes	$a = b = c$; $\alpha = \beta = \gamma = 90°$

Bravais calculated the unit area of a given lattice plane and showed that it is the inverse of the rectilinear area of the plane [28–36]. He ordered the lattice planes according to their decreasing rectilinear densities and developed the hypothesis that the most important planes of a crystal habit are those with maximum rectilinear densities. The Bravais law expresses this as "The cleavage planes of a crystal are the planes with the highest interplanar distances"; although as Fedorov and Friedel pointed out 50 years later, the order of the most important faces observed does not always coincide with the identical theoretical order because some faces may be missing. In modern terms crystal growth is controlled by kinetic as well as thermodynamic factors.

Weiss' major contribution was the classification of crystals into categories based on an analysis of the symmetry around the main axes. This regular arrangement defines what is now described as the crystal lattice. The classification of crystals as cubic, tetragonal, orthogonal, monoclinic and triclinic based on three axes and three angles and hexagonal and trigonal based on four axes three of which are separated by 120° which originate from this time are summarised in Table 2.

Fedorov had previously analysed the 2D wallpaper symmetry groups which can tile a Euclidian plane and demonstrated that they numbered 17. Bravais extended this to 3D lattices and showed that the additional lattice symmetries result from the translations of the lattice points involving combinations of the axes' directions. These result in the face-centred and body-centred lattices (the 14 Bravais lattices) illustrated in Fig. 4. The study and symmetry analysis of stereographic representations led Bravais to suggest that there were only 32 crystallographic groups permitted by symmetry. Sohncke, Fedorov, Schönflies and Barlow subsequently showed that the incorporation of additional translational symmetry operations, i.e. screw axes and glide planes, led to a total of 230 crystallographic space groups (see Fig. 5). Using the mathematics of group theory, Fedorov and Schönflies classified them in 1891. Barlow subsequently derived the space groups in an alternative fashion as part of his theoretical studies of close-packed spheres. Schönflies developed a notation for point groups, which is still used by theoreticians and spectroscopists, and extended it to three-dimensional crystallographic space groups by including the translational symmetry operations. It did not incorporate the specific symmetry

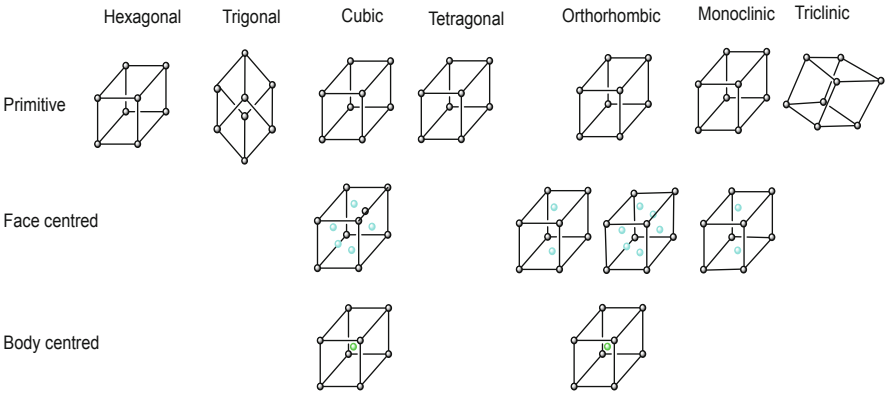


Fig. 4 The 14 three-dimensional Bravais lattices. For the face-centred and body-centred lattices, the additional atom lattice sites are shown in colour although all lattice points are equivalent

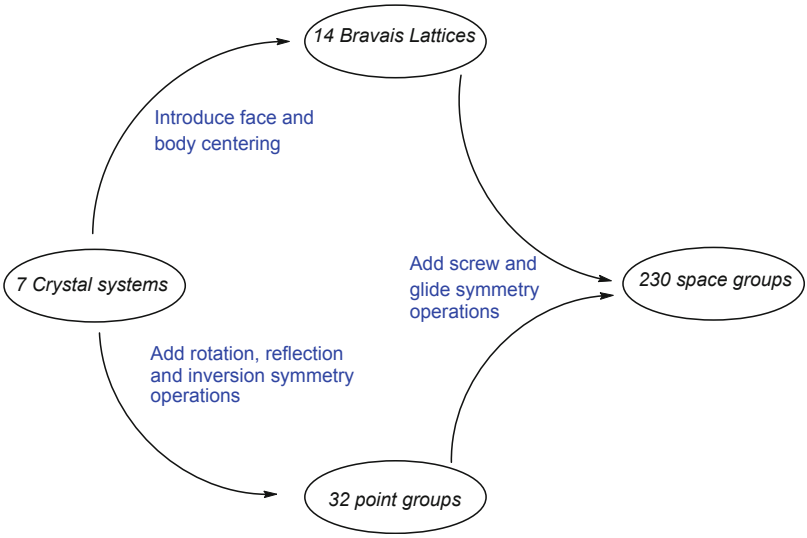


Fig. 5 The relationship between the 7 crystal systems, the 14 Bravais lattices, the 32 crystal point groups and the 230 space groups

operations for space groups, and the alternative Hermann-Mauguin notation was more widely adopted and published by the International Union of Crystallography. It proved to be an essential source book for crystallographers and was invaluable for the early crystallographic structural investigations when computers and the ability to access websites were not available. Model building and optical techniques were used to unravel the patterns observed in diffraction patterns, and these were greatly assisted by these symmetry considerations.

The idea of the close packing of atoms in crystals can be traced back to Kepler, and by the end of the nineteenth century, Barlow and Pope had developed its consequences for the periodic table. Barlow explored not only the packing of identical spheres but also combinations of spheres of different sizes, and his studies proved useful when the initial structures of simple salts were solved. Pope and Barlow's work interrelating crystal morphologies and crystal packing proved to be very helpful for W.L. Bragg and W.H. Bragg's analysis of the structures of alkali metal halides, zinc blende and calcite in the early days of X-ray crystallography [28, 33–36].

Finally, in order to emphasise the importance of symmetry in the first 50 years of X-ray diffraction, it is useful to digress to make a few comments concerning the proposal of the structure of DNA in 1952 [4]. The X-ray photographs on DNA taken in 1952 and particularly Rosalind Franklin's infamous photo 51 of hydrated β-DNA showed that it belonged to the monoclinic C2 space group and it had a diffraction pattern characteristic of a helical structure. The symmetry properties of the C2 space group implied that the DNA strand had a twofold symmetry axis. This piece of symmetry information implied an intertwined pair of helices proceeding in opposite directions. Model building based on putting together the base pairs through hydrogen bonds whilst maintaining the twofold symmetry axis contributed significantly to the proposed structure by Francis Crick and James Watson (Table 1 gives details of their Nobel Prize). The progress made in the subsequent 50 years takes us to the 2009 Nobel Prize which was awarded to Venkataraman Ramakrishna, Thomas Steitz and Ada Yonath for "studies of the structure and function of ribosome". A ribosome consists of the related RNA molecule and associated proteins and is found in the cytoplasm of living cells. Its biological function is to bind messenger RNA and transfer RNA to synthesise polypeptides and proteins. Yonath achieved the important tasks of producing the crystals, which are prerequisites of all diffraction experiments, and Steitz managed to solve the structure of these important molecules, and Ramakrishnan developed techniques which enabled the positions of single atoms to be identified (Table 1 gives details of their Nobel Prize).

4 Early Development of X-Ray Crystallography [19–27]

Von Laue's success depended more on intuition than careful analysis of the underlying physics because he had no equations to enable him to predict the possible appearance of the interference pattern [18, 21]. Based on traditional optics, he did not think that an interference pattern could result from white X-ray radiation, because of the wide range of frequencies generated by the X-ray source. He hoped that a crystalline sample with heavy atoms would produce a narrow band of fluorescence radiation which would produce an interference pattern via a secondary effect. William Lawrence Bragg the son of W.H. Bragg, a Professor of Physics at Leeds and who had passed onto him an interest in X-ray phenomena, recognised this flaw in his logic. The recent graduate in Natural Sciences at Cambridge speculated

that perhaps what was happening in the Laue experiment was that the different crystal planes which had geometric determined interplanar distances were selecting out from the white radiation those frequencies which gave the observed interference spots [19–23]. Moseley was more scathing in his criticism "The men who did the work (the German experiments of passing X-rays through crystals) failed to understand what it meant and gave an explanation which was obviously wrong. After much work Darwin and I found the real meaning of the experiments" [21, 40].

William Henry Bragg's interest in X-rays and γ-rays started round 1907, and he was sceptical of Stoke and Thomson's views that they were pulses of electromagnetic radiation and favoured their description as neutral pairs of material particles. Between 1885 and 1908, he was the Professor of Mathematics and Experimental Physics at the University of Adelaide, Australia, and returned to England to take up the chair at the University of Leeds. His son William Lawrence Bragg obtained a first-class honours degree in Mathematics in 1908 in Adelaide and returned to the England with his parents and enrolled at Trinity College Cambridge and obtained a first-class honours degree in Natural Science in 1912. The father had conveyed his enthusiasm for physics to his son, and they spent much time discussing his father's research and the recent developments in the study of X-rays. The combination of genetic inheritance, his father's interest in X-rays and a broad education in mathematics and the physical sciences meant that Bragg Jr. appreciated more clearly than the majority of 22 year olds the importance of von Laue's publication on the diffraction patterns of X-rays by crystals. In a remarkable example of synchronicity, he made a singular contribution to the development of science in the twentieth century. His interpretation of von Laue's X-ray diffraction patterns set in train a series of discoveries which enabled chemists to determine the structures of chemical compounds at the atomic level for the first time. This resulted in X-ray crystallography eventually becoming the most important physical technique for revealing the structures at atomic resolution not only for simple salts and molecules but also metals, alloys, ceramics, liquid crystals, hormones, vitamins and important classes of biologically important molecules, viz., proteins, enzymes and nucleic acids. Invariably, but not always, the structural determination resulted in a deeper understanding and interpretation of the function of the molecule. The Braggs were also instrumental in developing techniques which enabled the technique to be applied to increasingly complex species and ultimately to biological molecules. It was in W.L. Bragg's laboratories 40 years later, that Perutz and Kendrew solved the structures of myoglobin and haemoglobin and Watson and Crick proposed the structure of DNA. Many distinguished crystallographers received their training in X-ray crystallography under his guidance as postgraduate students or postdoctoral fellows. A more complete list of scientists who obtained Nobel Prizes for X-ray crystallography and related techniques are summarised in Table 1. With P.P. Ewald he was instrumental in setting up the International Union of Crystallography, which provided an important forum for progressing the interdisciplinary nature of the subject and the development of new techniques for obtaining X-ray crystallographic structural information for an extraordinary range of materials which went on to shape our century and provide the practitioners with the computer programmes and

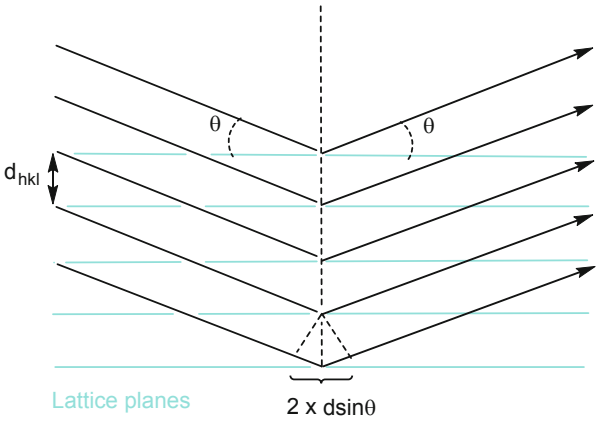


Fig. 6 The Bragg construction for diffraction by a three-dimensional crystal with one set of parallel lattice planes seen edge on. For a cubic crystal with cell dimension = a $d_{hkl} = a/(h^2 + k^2 + l^2)^{1/2}$

essential basic physical and mathematical procedures for accurately and reproducibly solving structures of ever increasing size and complexity.

Bragg Jr.'s undergraduate courses at Cambridge included theoretical and experimental optics given by C.T.R. Wilson [21]. Wilson's use of amplitude phase diagrams to illustrate diffraction and interference would prove to be a particularly important component of Bragg's breakthrough. He had been made aware that when white light fell on a diffraction grating, either it can be regarded as composed of all light of all colours which are sorted out by the grating into a spectrum or it can be regarded it as a regular pulse from which the grating manufactures a train of waves which are concentrated by interference effects. Whilst at Cambridge he also become familiar with the work of Pope and Barlow on the theory of crystal morphologies. As he started his research project with Professor J.J. Thomson on the theory of X-rays, he had a flash of inspiration. "As I was walking along the "Backs" by St John's College, I had a brainwave. X-ray diffraction could be considered in the same way as the diffraction of light by a diffraction grating, with the sheets of atoms playing the roles of the lines of a grating. As Authier has noted in his excellent book, *Early Days of X-ray Crystallography* [21], Bragg repeated Laue's calculation of the indices h_1, h_2, and h_3 of the spots on the published photograph using five different wavelengths for the X-rays, but he found that there were other sets of three integers which satisfied the interference requirement. Since this direction proved to be fruitless, he made the alternative assumption that the incident beam was composed of a wide range of wavelengths. If a pulse forms on a number of particles distributed over a plane, these particles act as centres of disturbance, and the secondary waves from them build up a wavefront, as if part of the pulse had been reflected from the plane as in Huygens' construction. Even if a minute part of the energy of a pulse is reflected by each plane in succession, this could lead to a significant interference maximum, because of the large number of parallel planes within the crystal [38–40]. The path length difference of pulses striking two successive parallel planes is $2d\sin\theta$ (as shown in Fig. 6), and the waves reinforce each other when $2d\sin\theta = n\lambda$, where n is an integer. The simplicity and elegance of this equation was immediately

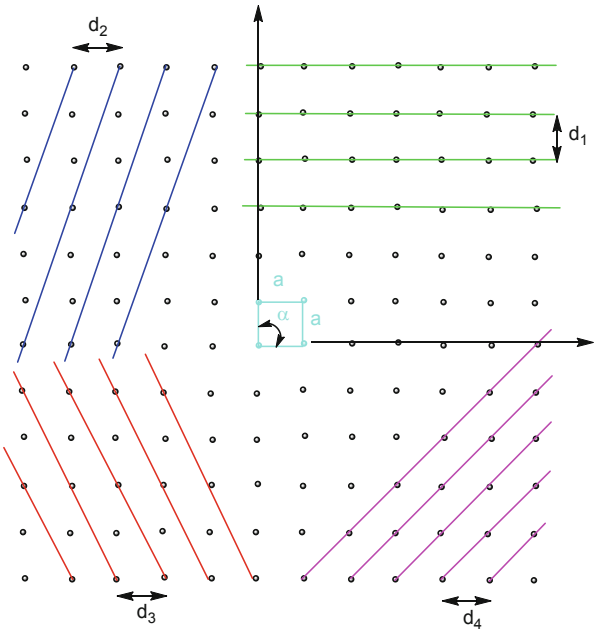


Fig. 7 Lattice planes projected onto the x,y plane for a unit cell with dimensions of a and $\alpha = 90^\circ$

recognised, and it has been described as Bragg's law ever since [19, 20, 23]. Bragg used the equation to reinterpret Laue's published results and specifically for zinc blende, ZnS, which Laue had assumed belonged to a simple cubic lattice. Barlow and Pope's previous optical work in developing the valence volume theory had shown it belonged to a face-centred cubic lattice. Bragg obtained a perfect fit between the observed and calculated diffraction spots for a face-centred cubic lattice, and this analysis provided a cell dimension for the lattice. Figure 7 illustrates schematically how different sets of planes with different d_{hkl} spacings could select different wavelengths to produce the observed spots in the observed diffraction pattern. His interpretation also accounted for the way in which the diffraction spots changed shape as the distance between the sample and the photographic plate was increased. As a 22-year-old research student, William Bragg presented his work to the Cambridge Philosophical Society in November 1912, and it was published in full on February 13, 1913, with a shorter version of his results having been published in December 2012 in *Nature*. He confirmed the reflective nature of the diffraction process by studying thin laminar samples of mica, which was known to form a layered structure which could be readily cleaved from more regular shaped crystals and confirmed that the observed spacings of the spots were consistent with the interference geometry illustrated in Fig. 6 and the individual spots satisfied the Bragg equation.

William Bragg Sr. and Jr. joined forces in the winter of 1912–1913 to explore more fully the possibilities of X-ray diffractometry. Bragg senior used his previous experience with X-rays tubes and ionisation detection chambers to design an X-ray

diffractometer, which resembled an optical spectrometer. The collimator was replaced by a lead block pierced by a hole, and the crystal was mounted like a prism on a rotating table, which could be rotated about a vertical axis. The X-rays were detected by an ionisation chamber mounted in such a way that it could rotate around the same vertical axis and connected to a Wilson gold-leaf electroscope. They estimated from their studies on rock salt and pyrite that the wavelength of the X-rays is 0.89×10^{-8} Å. They noted the sharpness of the diffraction peaks and surmised that the waves must occur in trains of great lengths and not in short pulses [21].

These experiments also established the field of X-ray spectroscopy and noted the relationships between the spectra of successive elements in the periodic table. This work was followed up by Moseley [41–43] who after working in Rutherford's Laboratory in Manchester returned to Oxford and showed using X-ray spectroscopy that the atomic number, i.e. the number of electrons and protons in a neutral atom, was the fundamental property of an element and not its atomic weight. This resulted because he established the relationship between the frequencies of the lines in X-ray spectra and the element's atomic number. Moseley showed that the square root of the frequencies of the corresponding lines in each spectrum increased by a constant amount when passing from one element to the next using the order of the Mendeleev periodic table except for nickel and cobalt. The X-ray spectrum of an element is entirely determined by one integer N which is equal to the charge on the nucleus, i.e. the atomic number of the element in the periodic table. This analysis led to Moseley to predict elements which had not been previously characterised, i.e. the elements technetium (43), promethium (61), hafnium (72) and rhenium (75). The Braggs' early work on crystallography recognised that different elements may diffract the X-rays differently and initially associated this with the atomic weight of the element, but subsequent results were consistent with Moseley's interpretation [21].

In that winter whilst Bragg Sr. was busy designing and supervising the construction of the X-ray spectrometer, Bragg Jr. worked with Pope and Barlow to familiarise himself more fully with how optical crystallographers used Miller indices to identify crystal faces and the definition of unit cells and space groups [21]. This knowledge of classical crystallography and the permission to borrow crystals from the Fitzwilliam Museum in Cambridge enabled the Braggs to solve the structures of several key compounds and establish the general usefulness of the technique. The early diffraction studies are summarised in Table 3 and underline how quickly the subject developed between 1913 and 1933. Bragg noted in his study of ZnS that the Miller indices of all the spots had indices which were either all even or all odd and thereby established the usefulness of systematic absences in defining the space group of a crystal. The use of systematic absences proved to be a very important feature of early structural determinations. This led to a confirmation that the ZnS structure was based on an infinite array of tetrahedra and led to a calculation of the ZnS distance. By 1920 the structures of 50 elements and compounds had been studied by this new technique although many of the major contributors had been diverted to assist their governments in war-related activities between 1914 and 1918. After the war the activity soon picked up, and by 1925 600 structures had been

Table 3 Early key structural determinations [25, 26]

Year	Description	
1913	Structure of cubic salts NaCl, KCl and ZnS	Bragg Jr
1913	Face-centred cubic diamond	Braggs
1914	CaF_2 (fluorite), FeS_2 (pyrite), $NaNO_3$, $CaCO_3$ (calcite), $Ca/MgCO_3$ (dolomite), $MnCO_3$ (rhodochrosite), $FeCO_3$ (siderite)	Braggs
1914	Copper metal	Bragg Jr
1914	Sulphur, quartz	
1915	$MgAl_2O_4$ (spinel), Fe_3O_4 (magnetite), Sb_2O_3 (senarmonite), NH_4Cl	Bragg Sr
1916	TiO_2 (rutile and anatase), SnO_2 (cassiterite), graphite (powder diffraction), ZrSiO4 (zircon), xenotime (YPO_4)	Vegard
1917	Graphite, iron, magnesium, and related metals (powder diffraction), SnO_2, chalcopyrite	Debye, Scherrer, Hull, Williams
1919	$Mn(OH)_2$ pyrochorite, $Mg(OH)_2$ brucite, $NaNO_3$, $CsICl_2$	Aminoff
1920	ZnO (zincite), ZnS – hexagonal (wurtzite), NiAs (nickeline), PbMoO4, CaWO4	Bragg Jr., Aminoff, Dickinson
1921	$[NH_4]_2[PtCl_6]$ Octahedral transition metal co-ordination	Wyckoff, Posnjak
1923	Hexamethylenetetramine $N_4[CH_2]_6$	Dickerson, Raymond
1928	Hexachlorobenzene C_6Cl_6	Lonsdale
1929–1933	Hexamethylbenzene C_6Me_6, anthracene, naphthalene	Lonsdale, Robertson

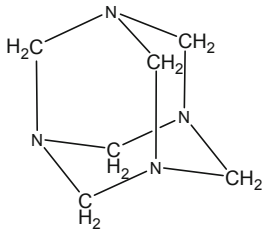
reported. Currently more than one million structures have been determined, and the instrumental and computing advances which have made this possible are discussed in subsequent chapters. As early as 1930, it was recognised that the increase in structural information was going to create archival and accession problems for future generations. R.W.G. Wyckoff made important contributions to addressing these issues and providing mechanisms for efficiently disseminating structural information. I shall return to this point later in subsequent chapter.

The Bragg's structural determinations of diamond which showed that all the carbon atoms had tetrahedral geometries confirmed the valency proposed initially by Kekulé and supported by van't Hoff and Lebel's geometric interpretation of optically active organic compounds [36, 37]. The consistency of tetrahedral geometries in organic compounds has stood the test of time. The octahedral geometries of the alkali metal halides derived from the diffraction experiments of the Braggs did not receive such universal acceptance and highlighted the necessity of defining both the co-ordination number and valency in inorganic compounds. The structures determined for CsCl and CsF showing the presence of eight coordinate metal ions required inorganic chemists to accept that alternative geometries were possible for ions, even in closely related compounds. The greater complexity of inorganic structural chemistry has remained a consistent pattern ever since and has led to a closer degree of collaboration between inorganic chemists and X-ray

crystallographers. The Braggs' elucidation of the structures of diamond and graphite some years later led to the first structural confirmation of allotropy. A phenomenon is recognised by Berzelius in the previous century [37]. Structural determinations in a number of laboratories in Europe, USA and Japan proceeded at a rapid rate until 1914, when the outbreak of the First World War required many young scientists to participate in their country's war efforts. Henry Moseley and William Henry Bragg's elder son sadly lost their lives whilst on active service. The structural studies of simple salts led to a better understanding of how the sizes of the atoms could influence the structures of salts and minerals. V.M. Goldschmidt was the first to tabulate covalent and ionic radii and begin to develop the structural principles of inorganic solid-state chemistry and lay the foundations of geochemistry [44].

In the 1916 papers, published independently by Kossel and Lewis, the fundamental descriptions of the chemical bond were defined [45–47]. They qualitatively described for the first time the dual extremes of chemical bonding by introducing ionic and covalent bonding types and suggested that the number of outer electrons of an atom played a very important role in determining the valencies and oxidation states of the elements [48]. Their ideas are still recognisable in current introductory undergraduate courses in chemistry. This duality was based on either the transfer or sharing of electrons to enable each atom to achieve an octet of electrons. Lewis' non-quantum mechanical description and his classical formulation in terms of localised electrons occupying the vertices of cubes reflected the influence of X-ray diffraction studies, and Kossel's description of electrons occupying circular orbitals owed much to Bohr's model for the hydrogen atom. Their descriptions gained credence from the X-ray crystallographic structures which were determined during this period.

The early structures utilised the classical knowledge of crystallography to assist the interpretation of the diffraction patterns, and the high symmetry of the crystals meant that the structures had few variable parameters. These early studies concentrated on well-formed crystals with high symmetries which could be solved readily once the cell dimensions and the symmetry of the unit cell were determined. Therefore, there was a concentration on inorganic crystals which established that the salts had very symmetric infinite structures. This gave strong support for the ionic description of chemical bonding proposed by Kossel. Support for Lewis' covalent bonding description was obtained in the structural determinations of graphite and diamond and those salts which had molecular anions, e.g. CO_3^{2-} and NO_3^- salts (see Table 3). The first structure of a simple organic molecule, hexamethylenetetramine, was completed in 1923 and showed it had the high-symmetry adamantane-related structure shown in Fig. 8. It crystallises in a body-centred cubic lattice which means that there are few variables associated with solving the structure and the high symmetry of the space group places each atom in a high-symmetry site. The lack of strong scattering associated with the hydrogen atoms meant that the tetrahedral geometries at carbon had to be implied rather than determined, and it was not until neutron diffraction was established after World War 2 that the positions of the hydrogen atoms were accurately located.



Hexamethylenetetramine

Fig. 8 Structure of hexamethylenetetramine

The first formal use of the space group classifications was made by Nishigawa in 1915 [49], and Niggli's "Geometrische Kristallographie des Diskontinuum" was widely used at this early stage of X-ray crystallography as an introductory guide [50]. In 1935 International Tables for the Solution of Crystal Structures edited by Hermann-Mauguin [51] and related publications IUCr proved to be invaluable for those entering the field. The crystals of common organic compounds provided arrangements of atoms, and their locations could be defined by X-ray diffraction which had profound and less easily surmountable problems because they belonged to primitive and monoclinic space groups, and consequently their diffraction patterns did not have the high symmetries observed for cubic and hexagonal crystals. The monoclinic space groups $P2_1/c$ and C2/c are statistically very prevalent in the structural chemistry of organic, organometallic and co-ordination chemistry.

The scattering abilities of carbon, hydrogen, oxygen and nitrogen, the common constituents of organic compounds, mean that they do not diffract as effectively as the atoms with higher atomic numbers. In addition, since each lattice point was associated with a molecule rather than an atom, the number of variables required to solve the structure increased rather rapidly as a function of the number of atoms. The structures of the first benzene derivatives were not reported until the late 1920s [52] (see Table 3). These showed that the C–C bonds in the benzene ring were all equal and shorter than those in diamond confirming Kekulé's model. The structure of $[NH_4]_2[PtCl_6]$ established the octahedral geometry around platinum, and it provided the first structural confirmation of Werner's theory of co-ordination compounds and encouraged the development of the dative bond notation in such compounds [53, 54].

5 Basic Physics

The basic physics underlying the scattering of X-rays was developed mainly by Darwin and Ewald around 1914 [55]. Initially Darwin established the mathematics associated with scattering by a single atom which was not affected by other neighbouring atoms and subsequently extended to scattering by successive planes of atoms which yielded results which Ewald had derived in his dynamic theory of

scattering. Qualitatively the scattering of X-rays results from the oscillation of the electrons in atoms which interact with the incoming electromagnetic waves of X-rays. Atoms with higher atomic numbers scatter more effectively because they have more electrons circulating the nucleus. The extent of scattering depends not only on the electron density distribution but also on the angle between the incoming X-rays and the diffraction plane, i.e. the scattering is larger in the forward direction and smaller in the reverse direction. The distribution of electron density in the plane involved in the Bragg reflection is expressed using a Fourier analysis which depends on the co-ordinates of the atoms and the scattering factors of the atoms located in that plane. These are summed over all the atoms in the unit cell, and the quantity is described as the structure factor F_{hkl}. The intensity of the diffraction spot is related to $|F_{hkl}|^2$. The Fourier methods were introduced in 1925 by W. Duane and R.J. Havighurst [56–57] who also modelled the thermal vibrations of the atoms. The measurement of the integrated intensities of the diffraction spots became possible either from the photographic data or by using more sophisticated X-ray measurements. Least squares methods were also developed in order to refine at the positional and thermal parameters and minimise the difference between the calculated and observed intensities. Initially the complexity of the calculations limited the number of variables, and it was not until the 1950s that computer programmes were developed to facilitate these calculations. Many of these programmes originated from the Oak Ridge National Laboratory, USA. The first 2D Fourier projections of the atomic densities were published in 1929 by Zachariason [58]. This procedure was extended by Lipson and Beevers [59] for the calculation of electron density difference maps in two dimensions based on Fourier analysis which proved particularly useful for solving the structures of organic molecules. Numerical methods were rapidly developed for the summation of Fourier syntheses optically by Bragg and mechanically using Beevers-Lipson strips.

The phase problem constitutes a fundamental limitation in X-ray crystallography which is ultimately related to the nature of measurement in quantum mechanics. X-ray detectors such as photographic plates or, in the modern age, charge-coupled devices (CCDs) measure only the intensity of the X-rays that hits them. This measurement is incomplete because an X-ray wave not only has an amplitude (related to the intensity) but also a phase, which is systematically lost in the measurement. In diffraction or microscopy experiments, the phase part of the wave often contains valuable information on the structure of the specimen. In X-ray crystallography, the diffraction data when properly assembled gives the amplitude of the 3D Fourier transform of the molecule's electron density in the unit cell. If the phases are known, the electron density can be calculated by Fourier syntheses [60, 61].

The Patterson method directly determines the positions of heavy atoms, which diffract the X-rays more effectively, in a structure [60, 61]. The Patterson function calculates the interatomic vectors and is dominated by those vectors between the heavy atoms. If the positions and scattering factors for these atoms are introduced into the Fourier calculation, because their phases dominate, then the resultant difference electron density map reveals other lighter atoms in the structure. The

Patterson method was ideal for studying the structures of organometallic and co-ordination compounds which were studied extensively after the discovery of ferrocene in 1953. In these compounds, the heavy metals dominate the calculated Patterson functions, and their co-ordinates in the unit cell can be calculated from the initial Patterson function. Insertion of these co-ordinates in the Fourier synthesis calculation leads to an initial set of phases for the electron density map. This set of phases is applied to the original amplitudes, and an improved electron density map is derived, which shows the positions of the lighter atoms. This process is repeated until an error term (usually R_{free}) has stabilised to a satisfactory value.

This technique was also widely used to determine the structures of organic and biologically important molecules. Dorothy Hodgkin's structural analysis of penicillin during World War 2 was made on metal salts of the penicillin anion. Hodgkin received the first penicillin crystals in autumn 1943 from the US pharmaceutical company Squibb, who had obtained high-quality crystals of benzylpenicillin (later known as penicillin G) as the sodium salt. Hodgkin made the potassium and rubidium salts so that she could compare the three different diffraction patterns. In the 1950s her determination of the structure of vitamin B_{12} was made possible because it is a co-ordination compound of cobalt which dominates the scattering. She also pioneered with Perutz and Kendrew the technique which introduced isomorphous heavy metal salts into protein crystals whilst retaining their isomorphous character in order to solve these complex structures. They were awarded Nobel Prizes in 1962 and 1964 for these contributions (see Table 1).

The 1953 monograph, "Solution of the Phase Problem I. The Centrosymmetric Crystal", by Karle and Hauptman contained the essential ideas necessary for solving the phase problem by direct methods [62, 63]. They were based on probabilistic methods and in particular the joint probability distributions of several structure factors. The direct method estimates the initial phases and expanding phases using a triple relation (A trio of reflections in which the intensity and phase of one reflection can be explained by the other two). A number of initial phases are tested and selected by this method. With the development of computers, the direct method became the most useful technique for solving the phase problem and is very widely used. In this monograph they also introduced the concepts of the structure invariants and semi-invariants, special linear combinations of the phases, and used them to devise recipes for origin specification in all the centrosymmetric space groups. The extension to the non-centrosymmetric space groups was made some years later. In the 1980s the problem of combining the traditional techniques of direct methods with isomorphous replacement and anomalous dispersion facilitated the solution of macromolecular crystal structures. For molecules whose crystals provide reflections in the sub-Ångström range, it is possible to determine phases by brute force methods, testing a series of phase values until spherical structures are observed in the resultant electron density map. This works because atoms have a characteristic structure when viewed in the sub-Ångström range. The technique is limited by processing power and data quality. For practical purposes, it is limited to "small molecules" because they consistently provide high-quality diffraction with very few reflections. In

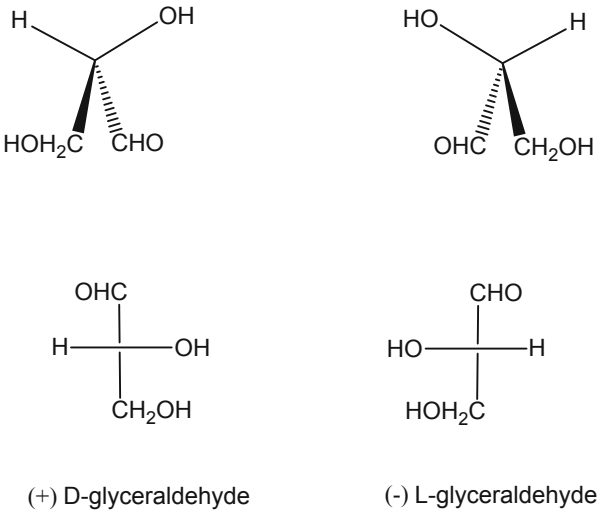


Fig. 9 The enantiomeric forms of the D- and L-glyceraldehyde enantiomers based on tetrahedral carbon were shown at the top and their Fisher projections below. The assignment of the D-formula to the compounds with the positive rotation was made by Fisher, and in 1951 the assignment was confirmed experimentally by Bijvoet using X-ray crystallography. The anomalous dispersion properties of the X-rays were used to determine the absolute configuration [64, 65]

1985 J. Karle and H.A. Hauptman were awarded the Nobel Prize for developing these direct methods for solving crystals structures.

Another powerful solution technique is the multi-wavelength anomalous dispersion (MAD) method [64]. In this technique, atoms' inner electrons absorb X-rays of particular wavelengths and reemit the X-rays after a delay, inducing a phase shift in all of the reflections, known as the anomalous dispersion effect. Analysis of this phase shift (which may be different for individual reflections) results in a solution for the phases. Since X-ray fluorescence techniques (like this one) require excitation at very specific wavelengths, it is necessary to use synchrotron radiation when using the MAD method. Other methods of experimental phase determination include multiple isomorphous replacement (MIR), where heavy atoms are inserted into structure (usually by synthesising proteins with analogues or by soaking), molecular replacement (MR) and single-wavelength anomalous dispersion (SAD).

The phenomenon of anomalous dispersion also had an important role in showing how X-ray crystallography could be used to determine the absolute configurations of optically active enantiomers of organic molecules. Compounds containing asymmetric carbon atoms have enantiomeric isomers of opposite chiralities which rotate polarised light in opposite directions. The enantiomeric forms of the D- and L-glyceraldehyde enantiomers based on tetrahedral carbon were shown at the top of Fig. 9 and their Fisher projections below. It was not possible to decide which of the two possible configurations corresponds to the isomer rotating the plane of light to the right. The assignment of the D-formula to the compounds with the positive rotation was made by Fisher, and in 1951 the assignment was confirmed experimentally by Bijvoet using X-ray crystallography. The anomalous dispersion properties of the X-rays were used to determine the absolute configuration.

Although the discussion above has emphasised the regular periodic nature of crystals developed from classical studies, these ideas played an important part in the

development X-ray crystallography. In recent years it has been discovered that it is also possible to form quasi-periodic crystals (quasicrystals), which have a structure that is ordered but not periodic. A quasicrystalline pattern can continuously fill all available space, but it lacks translational symmetry. According to the classical crystallographic restriction theorem, crystal lattices can only possess two-, three-, four-, and sixfold rotational symmetries. The Bragg diffraction pattern of quasicrystals shows sharp peaks with other symmetry orders, e.g. fivefold rotational symmetry. Aperiodic tilings were discovered by mathematicians in the early 1960s, and, some 20 years later, they were found to apply to the study of natural quasicrystals. The discovery of these aperiodic forms in nature has produced a paradigm shift in the fields of crystallography. In 1982 the materials' scientist Dan Shechtman observed that certain aluminium-manganese alloys produced the unusual diffractograms which today are seen as diagnostic of quasicrystalline structures. In 2009 after an extensive search, icosahedrite a naturally occurring mineral was shown to be a quasicrystal. Dan Shechtman was awarded the Nobel Prize in Chemistry in 2011 for his research in this area (see Table 1).

6 Spectacular Growth of Structural Data

By 1920 the structures of 50 elements and compounds had been studied by this new technique although many of the major contributors had been diverted to assist their governments in war-related activities between 1914 and 1918. After the war the activity soon picked up and by 1925 600 structures had been reported. Currently more than one million structures have been determined, and the instrumental and computing advances which have made this possible are discussed in subsequent chapters. As early as 1930, it was recognised that the increase in structural information was going to create archival and accession problems for future generations. R.W.G. Wyckoff made important contributions to addressing these issues and providing mechanisms for efficiently disseminating structural information [66].

These developments in theoretical chemistry provided the intellectual framework for chemistry over the last century and led to its exponential expansion into a wide range of new areas. The growth of chemistry can be appreciated by the following statistics. The Chemical Abstracts Service has estimated that it contains information on 100 million organic and inorganic compounds, and the *Cambridge Structural Database* (CSD) currently has crystallographic data on over a million inorganic and organic compounds. Only 4,500 structures had been solved and documented between 1923 and 1962 and formed the basis of the original database. Olga Kennard was the driving force behind setting up the original database. It is both a repository and a validated and curated resource for the three-dimensional structural data of molecules generally containing at least carbon and hydrogen. This means that it covers a very wide range of organic, metal-organic, organometallic and co-ordination molecules submitted by crystallographers and chemists from around the world. The database now contains more than a million entries. The specific

entries are complementary to the other crystallographic databases such as the Protein Data Bank (PDB), Inorganic Crystal Structure Database and the International Centre for Diffraction Data. Generally the entries are derived from X-ray crystallography, but it also has data based on electron and neutron diffraction. The other important databases for small molecules are the Crystallography Open Database and the Inorganic Crystal Structure Database. The Biological Macromolecular Crystallisation Database, the Protein Data Bank and the Nucleic Acid Database provide structural information on systems of biological importance.

A similar database has been developed for powder X-ray diffraction studies. The organisation was founded in 1941 as the Joint Committee on Powder Diffraction Standards (JCPDS). *The International Centre for Diffraction Data (ICDD)* was set up in 1978 and maintains a database of powder diffraction patterns. The Powder Diffraction File (PDF) includes the d-spacings and relative intensities of observable diffraction peaks for the samples. It has been designed to work with a diffractometer to identify unknown materials. With every entry, the database also contains bibliographic references, chemistry descriptions, structural classifications, crystallographic and physical properties. It also includes data on minerals and organic and pharmaceutical compounds. The 2019 release contains 893,400 unique data sets obtained from powder diffraction studies.

7 Related Diffraction Techniques

7.1 Introduction

Major contributions were made to the development of diffractometer and the conversion of the raw diffraction data into a model of the underlying atomic structure in the crystal, and within 50 years, X-ray crystallography had matured from solving the structures of simple inorganic salts to the determination of the structures of proteins, enzymes and nucleic acids (resulting in the development of molecular biology) and the elucidation of the structures of semiconductors, super-conductors, stereoregular polymers and new catalysts tor the interconversion of hydrocarbons (resulting in the subdisciplines of metallurgy and materials science) [5–14]. The discovery that the diffraction of X-rays could be used to unravel the secrets of crystals prompted physicists to explore the possibility that other atomic particles could generate diffraction patterns from inorganic crystals, liquids and gases. The contributions of Arthur Compton and Louis de Broglie (see Table 1) proved to be particularly important because they conclusively showed that all atomic particles have wave and particle properties and thereby completed Planck and Einstein's earlier work. Specifically, de Broglie proposed that the wavelength, λ, of an atomic particle with mass m and velocity v was connected by the following deceptively simple equation:

$$\lambda = h/mv$$

where h = Planck's constant.

The X-rays used in diffraction studies have a wavelength of ca 100 pm (Cu K_α: $\lambda = 154$ pm), and the calculated wavelength of 10 kV electrons is 12.2 pm and decreases to 2.5 pm for 200 kV electrons. Electron diffraction was confirmed 3 years later in 1927 by two independent experiments. At the University of Aberdeen, G.P. Thomson and his colleague A. Reid passed a beam of electrons through a thin film of celluloid and observed the predicted interference patterns. In the Western Electric Laboratories (subsequently the Bell Laboratories), C. J. Davisson and L. H. Germer guided their beam under vacuum conditions through crystalline nickel and obtained a diffraction pattern. In 1927 and 1929, Compton and de Broglie were awarded the Physics Nobel Prizes, and in 1937 Thomson and Davisson shared the Nobel Prize for Physics (see Table 1) for their experimental verification.

The development of the atomic bomb as part of the Manhattan project put the study of neutrons at centre stage, and neutron sources became more widely available to scientists through the commissioning of nuclear reactors. The possibility that neutrons may give diffraction patterns related to those which had been observed by Laue was investigated at the end of World War 2. The first neutron diffraction experiments were carried out in 1945 by Ernest Wollan using the graphite reactor at Oak Ridge National Laboratory, Tennessee, USA. He and Clifford Shull established the basic principles of the technique and applied it successfully to many different materials, e.g. the structure of ice and the microscopic arrangements of magnetic moments in materials. Shull was awarded the 1994 Nobel Prize and shared it with Bert Brockhouse for his development of neutron inelastic scattering (Table 1).

Since diffraction data can be obtained using X-rays, electrons and neutrons, it is useful to summarise the relative advantages and disadvantages of these three important structural techniques [67]. If the primary interest is to determine the positions of the atoms in a structure – X-ray diffraction is the preferred option because it is a relatively cheap technique; the diffractometers and X-ray sources are readily available from a range of international suppliers, and the computer programmes have ensured a rapid and seamless conversion of data into a description of the contents of the unit cell. The development of national synchrotron sources has created X-ray sources with high fluxes and with a small range of frequencies which makes it possible to study very small crystals which have large unit cells. Elements with low atomic numbers do not scatter the X-rays well, and consequently very light elements and especially hydrogen are difficult to locate accurately with X-rays [67, 68]. Structures containing atoms with large but similar atomic numbers can also create difficulties because it is difficult to distinguish atoms with similar structure factors, e.g. platinum and thallium [69], in large metal carbonyl clusters. Since X-rays are strongly attenuated as they pass through the walls of furnaces and cryostats, it proved to be more challenging to design equipment and to study the effects of temperature, pressure and other sources of electromagnetic radiation around the crystal target. Variable temperature and even high-pressure studies are now fairly routine since the

development of open flow cryostats (for temperature) and diamond anvil cells (for pressure).

Since electrons are negatively charged, they interact with both the electrons and nuclei of atoms in the diffraction process. The electron beam sources are relatively common, and the high fluxes mean that relatively small crystals can be studied using electron diffraction. It is particularly effective for studying thin samples. The wavelength can be readily tuned through the de Broglie expression, and the use of small wavelengths leads to more information from the diffraction experiment. Radiation damage of the crystals does occur and can be unacceptably high, and it is possible to reduce these effects by doing the experiments at very low temperatures. It is difficult to get useful information from electron diffraction studies of crystals which have magnetic centres because the electrons are also deflected by the local magnetic fields.

Neutrons have zero charge, and consequently furnaces and cryostats can be placed around the crystal target and not attenuate the neutron beam greatly. Once a neutron reactor was up and running, the costs of the neutrons and their costs and those for the adaptation of facilities are not prohibitive. These days, the majority of neutron studies are carried out at neutron spallation facilities (e.g. Oak Ridge and ISIS) where the neutrons can effectively be switched on and off and there is no longer the need to use a reactor in the old sense of the word. Neutrons have a magnetic dipole moment, and therefore they are ideal for probing magnetic structures [67]. In recent decades, they have been very useful for studying antiferromagnetic superstructures in transition metal oxides [70]. Neutron techniques have probed the vibrational, magnetic and lattice excitations (dynamics) of materials by measuring changes in the neutron momentum and energy simultaneously. The nuclear interactions between the neutrons and the nuclei in molecules are not large, and therefore relatively large crystals are required. It is particularly useful for studying compounds containing light atoms and in particular hydrogen and deuterium. This makes it suitable for studying organic polymers and biomolecules. The available fluxes are relatively low compared with those for X-rays, but their wavelengths can be changed over a wide range, and this means that they can be tuned to match the atomic spacings of particular interest in the crystal. The neutrons represent a relatively non-destructive probe, and the nuclear-neutron interactions are relatively easy to calculate.

7.2 Powder X-Ray Diffraction

Recording the diffraction patterns for single crystals using the Bragg condition was time-consuming because it required each diffraction spot to be measured individually. An obvious development was to study the effect of rotating the crystal or studying crystalline samples consisting of a very large number of small crystals randomly orientated, i.e. a finely ground powder. Powder X-ray diffraction was independently discovered by Paul Scherrer and Peter Debye in Göttingen, Germany,

and by Albert Hull at the General Electric Laboratories, Schenectady, USA [71–73]. In 1913 Debye had published papers which calculated the influence of lattice vibrations on the diffracted intensity (the Debye factor), and, in February 1915, he calculated the intensity of a diffracted beam for a random distribution of molecules. This encouraged him to propose to his student Paul Scherrer that he try and observe the diffraction of X-rays by a crystalline powder. They obtained very sharp lines when they studied the diffraction pattern from powdered LiF, and related studies on other halide salts confirmed the usefulness of the technique and also demonstrated its ability to show the presence of impurities. They also studied graphite and proposed that it belonged to a trigonal space group. In 1918 they published a very important analysis which calculated the effect of grain size on the line broadening of the powder diffraction patterns. The same year, Debye and Scherrer deduced from the analysis of the intensity of the diffraction lines that, in LiF, one valence electron is shifted from the lithium ion to the fluorine ion, a first step towards the study of electron density with X-ray diffraction. This area has attracted much interest in recent years because of the increased accuracy of the structural data and the ability to calculate the electron density distribution in molecules with great accuracy [74].

Hull [71, 72] did not have access to a single crystal of iron but discovered that he could obtain a good diffraction pattern from iron filings and interpreted the data on the basis of a body-centred cubic lattice. Observing the decrease in intensity with increasing Bragg angle, Hull concluded that the X-rays were diffracted by the electron cloud around the nucleus, and not by point diffraction centres. Diffraction by powders thus led both Debye and Scherrer and Hull to make observations of a fundamental nature. Hull then undertook an impressive series of 26 structure determinations of elements including Al, Ni, Li, Na and graphite. Debye and Scherrer had concluded that graphite was trigonal, but Hull's reinvestigation showed that it is hexagonal and based on layers of carbon atoms.

If a flat detector is placed perpendicular to the incident beam, on the opposite side of the sample, the set of reflections from the multiple crystals appear as continuous circles. The overall result is a set of concentric circles with radii governed by the Bragg equation. Therefore, the information obtained is identical to that gained from the conventional Bragg experiment, i.e. the unit cell dimensions and unit cell symmetry are in principle identical. The net effect of using powder sample rather than a single crystal is to effectively compress the information into a one-dimensional output where the only variable is θ. This technique proved to be particularly useful for studying minerals and solid-state inorganic compounds which crystallised in high-symmetry space groups. For high-symmetry crystals, it is possible to assign each circle a unique Miller index and measure its intensity. The unit cell structure can then be solved from these data [75–77]. The compression of the data in this way can be problematic when the diffraction circles for different (*h,k,l*) reflections have similar values of θ. This is particularly problematic for monoclinic and triclinic crystals. The technique proved to be very useful for studying interrelated isomorphous compounds with the same stoichiometry and was very effectively developed by Goldschmidt to study minerals and solid-state compounds [44]. Within a few years, this group had solved the structures of more than 200 structures

containing 70 elements and used the structural data to define the characteristic ionic and covalent radii of the atoms in the compounds. For the synthetic solid-state chemists and physicists, the technique proved to be invaluable for determining whether the powder was pure. The impurity could also be identified if its diffraction data had been recorded previously, and its relative concentration could also be established from the intensities of the lines [75–77].

It also enabled chemists to rapidly measure how the cell dimensions and perhaps the phase of the crystal changed as a function of temperature. Such studies were of great importance in the development of metallurgy during the twentieth century. Powder diffraction was quickly adopted because it did not require specimens which formed large crystals, and many samples could be studied rapidly. It was widely applied in mineralogy, petrology, metallurgy and materials science. One of the difficulties of the method lies in the indexation of the diffraction lines for samples which crystallised in lower-symmetry space groups. The determination of more complex structures had to wait for the development of new refinement techniques [78, 79]. This led to a complete renewal of the powder diffraction method, with applications to the study of different classes of new materials in chemistry, materials science and biology. The X-ray powder diffraction file now contains nearly a million entries. In recent years increased computer speeds and more sophisticated programmes have made it possible to use the technique for more complicated organic and organometallic compounds.

The diffraction lines in a powder diagram are also very sensitive to the degree of perfection of the crystal. A general analysis, taking distortions within the grains into account, was given by A. R. Stokes and A. J. C. Wilson [80, 81]. The profile of the lines depends both on the size of grains, as mentioned above, and on the distribution of defects. B. E. Warren and his school studied in the 1950s the influence of online shape of microtwins and stacking faults such as those introduced during cold working of metals and alloys. During annealing of these materials, the grains recrystallise along preferred orientations, their size increases, and the diffraction lines become discontinuous [82, 83].

7.3 Neutron Diffraction

The de Broglie equation established that neutrons generated in a nuclear reactor (or more recently a spallation source) had wave lengths in the same range as that established for X-rays and consequently could also be diffracted by crystals. The high initial cost of neutron sources and the diffracted intensities are rather weaker than those obtained with X-rays because the diffraction centres are not the electrons but the nuclei. Neutron diffraction has proved to particularly useful for locating hydrogen atoms in structures because the hydrogen atom has only a single electron and it is not a strong diffraction centre [84–88]. Since the nuclei and the neutron are both small, significant scattering only occurs when the neutron is close to the nucleus. The advantage is that the scattering does not fall off with Bragg angle as

steeply as X-rays. There is no simple relationship linking the scattering factors of atoms and their place in the periodic table. This makes it necessary to use larger crystals than those commonly used in X-ray diffractions studies, but nonetheless it has proved invaluable for determining the geometries of metal hydrides' agostic interactions involving C–H bonds and transition metals, the location of bridging hydrides in boranes and carboranes and molecules exhibiting hydrogen bonding interactions [85, 86]. Although the scattering power of an atom is directly related to the number of electrons in the neutral atom, there is no simple relationship between the neutron scattering power and the atomic number. The neutron scattering powers vary erratically, and there may be large differences between adjacent atoms, and different isotopes of the same atom may show different scattering powers. The phase of the scattering may have positive and negative signs, for example, D has a positive scattering factor of 6.7, whereas for hydrogen it is −3.7. To generalise neutron diffraction is very useful for locating the positions of light atoms more accurately, e.g. the location of H and D in partially deuterated compounds. The technique has proved to be particularly important for studying and defining the nature of the interactions in hydrogen-bonded systems [89–94]. For such studies the bond length involving hydrogen determined from neutron data will differ from that derived from X-ray data because the former locates the nucleus, whereas the latter locates the electron density which is not completely localised on the hydrogen but shifted towards the atom to which it is bonded as a result of covalency effects. For example, a C–H bond which has a length of 1.08 Å in a neutron study would have length of 0.98 Å by X-ray diffraction.

More generally the careful determination of a structure using X-ray and neutron data may be used to more carefully map the electron density and lead to conclusions about bonding effects. However, it must be noted that if the valence electrons represent only a small amount of the total number of electrons, then these studies have to be completed with great care [90–95]. The neutrons define accurately the positions of the nuclei, and the X-ray data provides information concerning the distribution of electron density throughout the molecule and consequently provides an insight on how the electron density has been modified as a result of ionic and covalent bonding effects. This area which was pioneered by P. Coppens [95, 96] and many of these studies have used the R. Bader's quantum mechanical analysis of the topology of the electron density function for interpreting the three-dimensional space electron density [97, 98].

Although neutrons are uncharged, they carry a magnetic moment and therefore interact with magnetic moments associated with the target crystal, including those arising from the electron cloud around an atom. Neutron diffraction can therefore reveal the microscopic magnetic structure of a material. Magnetic scattering does require an atomic form factor as it is caused by the much larger electron cloud around the tiny nucleus. The intensity of the magnetic contribution to the diffraction peaks will therefore decrease towards higher angles [99].

One major advantage of neutron diffraction over X-ray diffraction is that the latter is rather insensitive to the presence of hydrogen (H) in a structure, whereas the nuclei 1H and 2H (i.e. Deuterium, D) are strong scatterers for neutrons. The greater

scattering power of protons and deuterons means that the position of hydrogen in a crystal and its thermal motions can be determined with greater precision by neutron diffraction. The structures of metal hydride complexes, e.g., $Mg_{2[}FeH_6]$ and $K_2[ReH_9]$, have been assessed by neutron diffraction [86, 95]. The neutron scattering lengths bH = −3.7406(11) fm [8] and bD = 6.671(4) fm, [13a] for H and D, respectively, have an opposite sign, which allows the technique to distinguish them. In fact there is a particular isotope ratio for which the contribution of the element would cancel; this is called null-scattering.

It is undesirable to work with the relatively high concentration of H in a sample. The scattering intensity by H-nuclei has a large inelastic component, which creates a large continuous background that is more or less independent of scattering angle. The elastic pattern typically consists of sharp Bragg reflections if the sample is crystalline. They tend to drown in the inelastic background. This is even more serious when the technique is used for the study of liquid structure. Nevertheless, by preparing samples with different isotope ratios, it is possible to vary the scattering contrast enough to highlight one element in an otherwise complicated structure. The variation of other elements is possible but usually rather expensive. Hydrogen is inexpensive and particularly interesting, because it plays an exceptionally large role in biochemical structures and is difficult to study structurally in other ways.

7.4 Electron Diffraction

It was noted earlier that in the early days the determination of the structures of molecular organic and organic compounds by X-ray diffraction proved problematic because of the phase problem. Electron diffraction proved to be very important between 1920 and 1960 for establishing the structures of small organic and inorganic molecules [100]. The only constraint was that the molecule had to be sufficiently volatile to generate a good concentration of molecules in the gas phase. When a monochromatic beam of electrons is used with molecules in the gas phase it interacts only at the point where the beam crosses. The electrons are scattered mainly by interactions with the electric fields of the atomic nuclei contained in the molecule, and the proportion of electrons scattered is large (larger than those for X-rays and much larger than those for neutrons [101–103]. The data are collected within a matter of seconds or at most minutes. However, the intensity of the scattering falls off rapidly with the scattering angle. In a gaseous state, the molecules are randomly orientated, and therefore the diffraction pattern consists of diffuse concentric rings and is described simply of intensities as a function of scattering angle. For a polyatomic molecule, the total molecular scattering is the sum of the components of each pair of atoms on the molecule, i.e. for a triatomic molecule there are three components, 6 for 4 atoms, 10 for 5 atoms and 1,225 for 50 atoms. Thus, only relatively simple molecules can be completely structurally characterised by electron diffraction in the gas phase. The data is more amenable to analysis if the molecules are highly symmetric. It is possible to combine information obtained from other

sources, such as rotational spectra, NMR spectroscopy or high-quality quantum mechanical calculations with the electron diffraction data, if the latter are not sufficient to determine the molecule's structure completely. This makes it ideal for studying small molecules, and Pauling and his colleagues at the California Institute of Technology used the technique very extensively in the 1930s and 1940s to enlarge the chemists' structural database. References [104–116] provide examples of typical applications.

Organic molecules studied include acrolein, glyoxal, butadiene, ethylene, 1,3,5-cis-hexatriene, 1,3,5-trans-hexatriene, 2,3-dimethylbutadiene and 3,4-dimethyl-1-2,4-hexadienes. Some systems involving triple bonds which were also studied are acrylonitrile, vinyl acetylene, cyanogen, diacetylene and dimethylacetylene. Strained ring compounds studied include norbornadiene, 1,4-dichloronorbornane, bicyclo-[1,1, l]-pentane, bicyclo-[2.1.0] pentane, spiropentane, bicyclo-[3.1.1]-heptane, bicyclo-[2.1.1] hexane, 4-chloronortricyclene, bullvalene, hexamethyl Dewar benzene, hexafluoro-Dewar benzene, bicyclo-[2.2.2]-octane and triethylenediamine. It proved to be particularly useful for defining changes in dimensions and bond angles for isostructural molecules. For example, the C–C bond length in C_2F_4, C_2H_4, C_2Cl_4 and C_2Br_4 were shown to correlate with the stretching force constant. Figure 10 gives examples of simple inorganic molecules which were studied, and the bond length differences have been discussed by Haaland [107].

In the area of organometallic chemistry [109, 110], cyclopentadienyl compounds of the main group and transition metals have been extensively studied. In iron and ruthenium dicyclopentadienyl, the rings appear to be eclipsed (D_{5h}), but the barrier to internal rotation is only 1 kcal/mol. In dicyclopentadienylnickel and dicyclopentadienylmanganese, essentially free rotation was observed. In lead and tin cyclopentadienyl, the metal atom is so bulky that the rings are no longer parallel and are tilted relative to each other by 45 and 55°, respectively. Bartell was instrumental in testing the ideas of Nyholm and Gillespie by determining numerous structures of central atom molecules [114–116]. The structures of a number of central atom molecules with three, five, six and seven substituent groups have been investigated by electron diffraction, and the results have proven to be useful in testing various bonding theories. It is not essential to undertake the studies at ambient temperatures, and high-temperature species have been studied, for example, As_4, AsI_3 and GaI_3. The ED patterns for all of the dihalides of Be, Mg, Ca, Sr, Ba, Zn, Cd and Hg with the exception of HgF_2 have been obtained and the bonded distances and angles determined. The angular nature of the heavier Group 2 metal halides has resulted in extensive theoretical speculations [107, 108].

Low-energy electron diffraction (LEED) is a technique for the determination of the surface structure of single-crystalline materials by bombardment with a collimated beam of low-energy electrons (20–200 eV) and observation of diffracted electrons as spots on a fluorescent screen.

LEED may be used in one of two ways:

1. Qualitatively, where the diffraction pattern is recorded and analysis of the spot positions gives information on the symmetry of the surface structure. In the

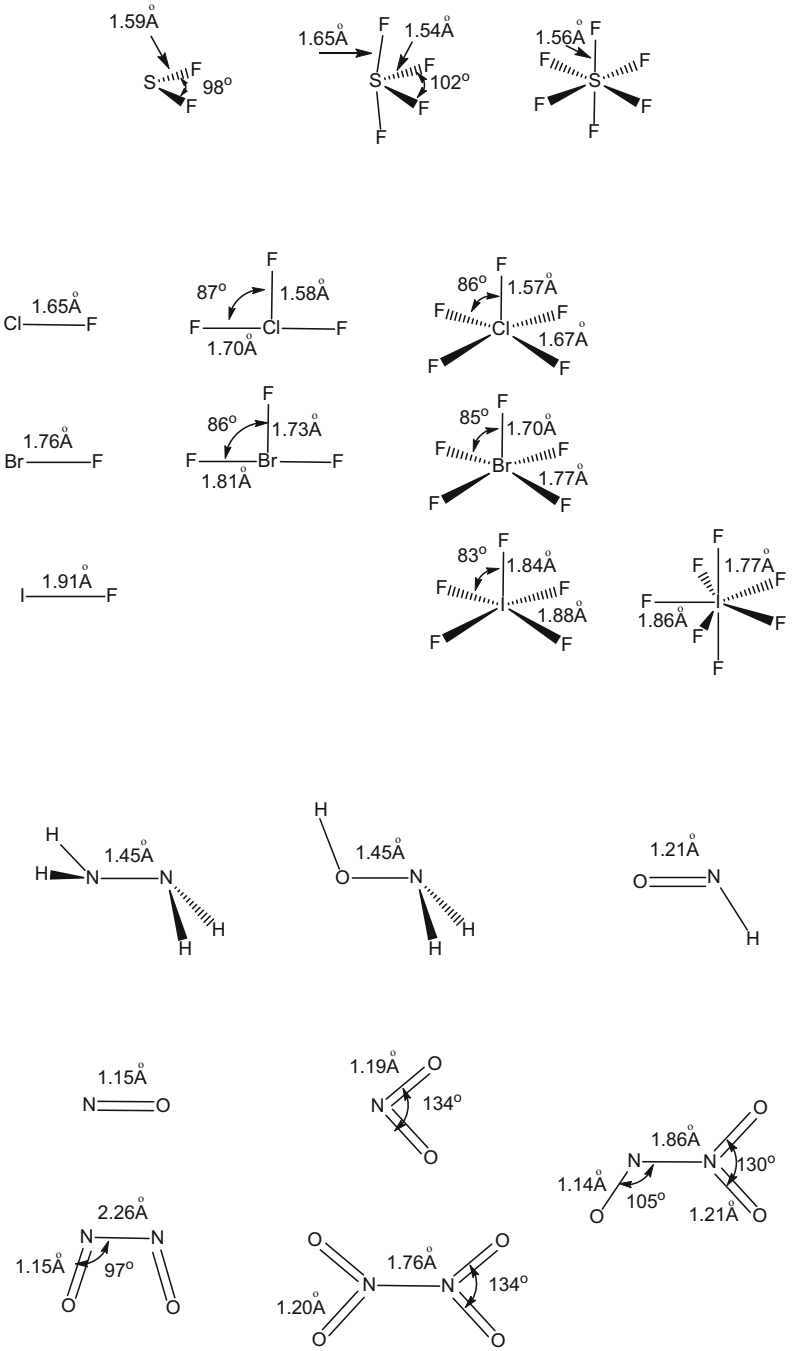


Fig. 10 Examples of main group fluorides and oxides studied by electron diffraction techniques in the gas phase

presence of an adsorbate, the qualitative analysis may reveal information about the size and rotational alignment of the adsorbate unit cell with respect to the substrate unit cell.
2. Quantitatively, where the intensities of diffracted beams are recorded as a function of incident electron beam energy to generate the so-called I-V curves. By comparison with theoretical curves, these may provide accurate information on atomic positions on the surface at hand.

Reflection high-energy electron diffraction (RHEED) is a technique used to characterise the surface of crystalline materials. RHEED systems gather information only from the surface layer of the sample, which distinguishes RHEED from other characterisation methods for materials that also rely on diffraction of high-energy electrons.

7.5 Electron Microscopy

An electron microscope is a microscope that uses a beam of accelerated electrons as a source of illumination. As the wavelength of an electron can be up to 100,000 times shorter than that of visible light photons, electron microscopes have a higher resolving power than light microscopes and can reveal the structure of smaller objects but not quite at the atomic scale. Electron microscopes use shaped magnetic fields to form electron optical lens systems that are analogous to the glass lenses of an optical light microscope. Electron microscopes are used to investigate the ultrastructure of a wide range of biological and inorganic specimens including microorganisms, cells, large molecules, biopsy samples, metals and crystals. Industrially, electron microscopes are often used for quality control and failure analysis. The technique was pioneered in 1933 by Ernst Ruska, and he was awarded a Nobel Prize in 1986. Modern electron microscopes produce electron micrographs using specialised digital cameras and frame grabbers to capture the images. The transmission electron microscope (TEM) uses a high-voltage electron beam to illuminate the specimen and create an image. The electron beam is produced by an electron gun, commonly fitted with a tungsten filament cathode as the electron source. The electron beam is accelerated by an anode typically at +100 keV (40–400 keV) with respect to the cathode, focused by electrostatic and electromagnetic lenses and transmitted through the specimen that is in part transparent to electrons and in part scatters them out of the beam. When it emerges from the specimen, the electron beam carries information about the structure of the specimen that is magnified by the objective lens system of the microscope. The spatial variation in this information (the "image") may be viewed by projecting the magnified electron image onto a fluorescent viewing screen coated with a phosphor or scintillator material such as zinc sulphide. Alternatively, the image can be photographically recorded by exposing a photographic film or plate directly to the electron beam, or a high-resolution phosphor may be coupled by means of a lens optical system or a fibre optic light

guide to the sensor of a digital camera. The image detected by the digital camera may be displayed on a monitor or computer. A scanning transmission electron microscope has achieved better than 50 pm resolution in annular dark-field imaging mode and magnifications of up to about 10,000,000×, whereas most light microscopes are limited by diffraction to about 200 nm resolution and useful magnifications below 2000×.

Molecular biologists have attempted to combine these techniques in order to solve important problems and Klug used methods from X-ray diffraction, microscopy and structural modelling to develop crystallographic electron microscopy in which a sequence of two-dimensional images of crystals taken from different angles are combined to produce three-dimensional images of the target. He studied the structure of transfer RNA and found what is known as zinc fingers as well as the neurofibrils in Alzheimer's disease. He received the 1975 Nobel Prize for his research in this area. This approach has progressed in the subsequent decades, and Joachim Frank, Jacques Dubochet and Richard Henderson were awarded the Nobel Prize in 2017 for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules (Table 1).

Between 1975 and 1986, Joachim Frank developed an image processing method in which the electron microscopes' fuzzy two-dimensional images are analysed and merged to reveal a sharp three-dimensional structure. Jacques Dubochet discovered how water could be retained in the electron microscopy sample chamber. Liquid water evaporates in the electron microscope's vacuum, which makes the biomolecules collapse. In the early 1980s, Dubochet succeeded in vitrifying water – he cooled water so rapidly that it solidified in its liquid form around a biological sample, allowing the biomolecules to retain their natural shape even in a vacuum. In 1990 Richard Henderson succeeded in using an electron microscope to generate a three-dimensional image of a protein at atomic resolution. This breakthrough proved the technology's potential. The desired atomic resolution was reached in 2013, and researchers can now routinely produce three-dimensional structures of biomolecules. In the past few years, scientific literature has been filled with images of everything from proteins that cause antibiotic resistance to the surface of the Zika virus. Biochemistry is now facing an explosive development and is all set for an exciting future [117].

8 Summary

In this chapter, I have attempted to give a qualitative introduction to the early history of X-ray crystallography and a summary of the related diffraction techniques which have been developed subsequently. In this way, I hope that I have provided a suitable introduction to the subsequent chapters which address the current state-of-the-art issues in studying the structures of small molecules. For those who wish to delve deeper into the mathematics and technical developments, I can strongly recommend references [5–14]. References [19–27] provide more details concerning the early

historical development of the subject and a deeper appreciation of the important personalities. There is a second chapter which will summarise the way in which small molecule crystallography has been a major influence on the way bonding theories have emerged from the wealth of structural data which the technique has made available. Other chapters describe the way in which the vast amount of structural information which is being generated is archived and the development of new techniques which have enhanced the subject.

References

1. Mingos DMP (ed) (2016) 50 years of structure and bonding – the anniversary volume. Structure and bonding, vol 172, pp 1–374
2. Stalke D (ed) (2012) Electron density and chemical bonding 1. Structure and bonding, vol 146, pp 1–207
3. Stalke D (ed) (2012) Electron density and chemical bonding I1. Structure and bonding, vol 147, pp 1–229
4. Anders Lilja A, Liljas L, Ash M-R, Lindblom G, Nissen P, Kjeldgaard M (2017) Textbook of structural biology, vol 8, 2nd edn. World Scientific Publishing, London
5. Giacovazzo C, Manaco HL, Vierbo D, Scordari F, Gilli G, Catti M (1992) Fundamentals of crystallography. International Union of Crystallography, Oxford University Press, Oxford
6. Clegg W (1998) Crystal structure determination. Oxford University Primers, Oxford University Press, Oxford
7. Sands DE (1975) Introduction to crystallography. Dover Publications, New York
8. Rhodes G (2006) Crystallography made crystal clear, 3rd edn. Elsevier, Oxford
9. Stout GH, Jensen LH (1989) X-ray structure determination, 2nd edn. Wiley, New York
10. Luger P (1980) Modern X-ray analysis on single crystals. de Gruyter, Berlin
11. Ladd MFC, Palmer RA (1985) Structure determination by X-ray crystallography. Plenum, New York
12. Glusker JP, Trueblood KN (1985) Crystal structure analysis, 2nd edn. Oxford University Press, Oxford
13. Hammond C (1997) The basics of crystallography and diffraction. Oxford University Press, Oxford
14. Wolfson MM (1997) An introduction to X-ray crystallography. Cambridge University Press, Cambridge
15. Röntgen WC (1895) Uber eine neu Art von Strahlen. Sitzungber. Der Wurzburger Physik-Medic Gesekksch 137:132–141. Translated into English by Stanton A (1896) On a new kind of rays. Nature 53:274–276
16. Thomson JJ (1897) Cathode rays. Philos Mag 44:293–303
17. Friedrich W, Knipping P, von Laue M (1912) Sitzungsberichte der Math. Phys. Klasse (Kgl.) Bayerische Akademie der Wissenschaften, pp 303–322
18. von Laue M (1913) Kritische Bemerkeungen zu den Deutungen der Photogramme von Freidrich und Knipping. Phys Z 14:421–423
19. Bragg WL (1962) Ewald PP (ed) Fifty years of X-ray diffraction. International Union of Crystallography, Oxford University Press, Oxford, pp 531–539
20. Thomas JM, Phillips DC (eds) (1990) Selections and reflections: the legacy of Sir Lawrence Bragg. Science Reviews, London
21. Authier A (2013) Early days of X-ray crystallography. International Union of Crystallography, Oxford University Press, Oxford
22. Kemp TJ, Alcock NW (2017) 110 years of X-ray crystallography. Sci Prog 100:25–44

23. Thomas JM (2012) Centenary – the birth of X-ray crystallography. Nature 491:186–124
24. Thomas JM (2012) WL Bragg – the pioneer of X-ray crystallography and his pervasive influence. Angew Chem Int Ed 51:12946–12958
25. Bijvoet JM, Burgers WG, Hagg G (eds) (1969) Early papers on diffraction of X-rays by crystals I. International Union of Crystallography, A. Oosthoek's Uitgevrsmaatschappij, Utrecht
26. Bijvoet JM, Burgers WG, Hagg G (eds) (1972) Early papers on diffraction of X-rays by crystals II. International Union of Crystallography, A. Oosthoek's Uitgevrsmaatschappij, Utrecht
27. Bragg WL, Phillips DC, Lipson H (1992) The development of X-ray analysis. Dover, New York
28. Wooster WA (1990) A brief history of physical crystallography. Kluwer, Dordrecht
29. Bloss FD (2002) Optical microscopy, Mineralogical Society of America Monograph Series, no. 5. USA
30. Kraus EH, Hunt WF, Ramsdell LS (1959) An introduction to the study of minerals and crystals, 5th edn. McGraw-Hill Books, New York
31. Wood EA (1964) Crystals and light – an introduction to optical crystallography. van Nostrand, New York
32. Wood EA (1963) Crystal orientation manual. Columbia University Press, New York
33. Glazer AF (2016) Crystallography. A very short introduction. Oxford University Press, Oxford
34. Glusker JP (1990) A brief history of chemical crystallography. Kluwer, Dordrecht
35. Wells AF (1956) The third dimension in chemistry. Clarendon Press, Oxford University Press, Oxford
36. Wells AF (1984) Structural inorganic chemistry, 5th edn. Clarendon Press, Oxford University Press, Oxford
37. Mingos DMP (2019) The periodic table 1. Struct Bond 181:1–50
38. Bragg WL (1912) The specular reflection of X-rays. Nature 90:410
39. Bragg WL (1913) The diffraction of short electromagnetic waves by a crystal. Math Proc Cambridge Philos Soc 17:43–57
40. Bragg WL (1913) The structure of some crystals as indicated by their diffraction of X-rays. Proc R Soc Lond A89:248–277
41. Moseley HGF, Darwin CG (1913) The reflection of X-rays. Nature 90:594–594
42. Moseley HGJ (1913) The high frequency spectra of the elements. Philos Mag 26:1024–1034
43. Moseley HGJ (1914) The high frequency spectra of the elements II. Philos Mag Ser 6 27:703–713
44. Goldschmidt VM (1926) Geochemische Verteilungsgesetze der Elemente. In: Skrifter Norske Videnskaps – Akad. Oslo, (I) Mat. Natur. This is an 8 volume set of books
45. Lewis GN (1916) The atom and the molecule. J Am Chem Soc 38:762–785
46. Lewis GN (1923) Valence and the structures of atoms and molecules. The Chemical Catalog, New York
47. Kossel W (1916) Formation of molecules and its dependence on atomic structure. Ann Phys 49:229–362
48. Mingos DMP (2016) The chemical bond 1. Struct Bond 169:1–251
49. Nishigawa S (1915) Structure of some crystals of spinel group. Proc Tokyo Math Phys Soc 8:199–209
50. Niggli P (1919) Geometrische Kristallographie des Diskontinuums. Gebrüder-Bornträger, Leipzig
51. Sharma BD (1982) Crystallographic and spectroscopic symmetry notations. J Chem Educ 59:554–557. gives a good discussion of these alternatives
52. Lonsdale K (1928) The structure of the benzene ring. Nature 122:810–810
53. Sidgwick NV (1923) The nature of the non-polar link. Trans Farad Soc 19:469–475

54. Constable EC, Housecroft CE (2013) Coordination chemistry – the legacy of Alfred Werner. Chem Soc Rev 42:1429–1439
55. Darwin CG (1914) The theory of X-ray reflection. Philos Mag Ser 27(6):315–333; 675–690
56. Duane W (1925) The calculation of X-ray diffracting power at points in a crystal. Proc Natl Acad Sci 11:489–493
57. Havighurst RJ (1925) The distribution of diffracting power in sodium chloride. Proc Natl Acad Sci 11:502–507
58. Zachariasen WH (1929) The crystal structure of potassium chlorate Z. Kristallografiya 71:501–506
59. Lipson H, Beevers CA (1936) An improved method of two-dimensional Fourier syntheses for crystals. Proc Phys Soc A48:772–780
60. Patterson A (1934) A Sourier series method for the determination of the components of the interatomic distances in crystals. Phys Rev 46:372–376
61. Patterson A (1935) A direct method for the determination of the components of the interatomic distances in crystals Z. Kristallografiya 90:517–542
62. Hauptman HA (1990) History of X-ray crystallography. Struct Chem 6:617–620
63. Hauptman HA, Karle J (1953) The solution of the phase problem I. The centrosymmetric crystal: American Crystallographic Association monograph no 3. Polycrystal Book Service, Dayton
64. Bijvoet JM, Peerdeman JM, Van Bommel AJ (1951) Determination of the absolute configuration of optically active compounds by X-rays. Nature 168:271–273
65. Monaco HL (1992) Experimental methods in X-ray crystallography. In: Giacovazzo C, Manaco HL, Vierbo D, Scordari F, Gilli G, Catti M (eds) Fundamentals of crystallography, international union of crystallography. Oxford University Press, Oxford, pp 229–318
66. Wyckoff RWG (1935–1971) Crystal structures, vols 1–6. Wiley, New York
67. Ebsworth EAV, Rankin DWH, Cradock S (1991) Structural methods in inorganic chemistry, 2nd edn. Blackwell Scientific Publications, London
68. Welch AJ (2017) What can we learn from the structures of metallocarboranes? Crystals 7:234
69. Mednikov EG, Dahl LF (2013) How innocent is thallium(I)? Corrected formula of clusters previously reported as Au_2Pd_{14} and $AuPd_9$ carbonyl clusters. Chem Commun 49:1085–1087
70. West AR (2000) Basic solid state chemistry, 2nd edn. Wiley, Chichester, p 162
71. Hull AW (1917) The crystal structure of iron. Phys Rev 9:84–90
72. Hull AW (1917) The crystal structure of magnesium. Proc Natl Acad Sci 3:470–473
73. Debye P, Scherrer P (1916) Interferenz an regelos orintierten Teilschen im Röntgenlicht 1. Phys Z 17:277–285
74. Coppens P (2017) The dramatic development of X-ray crystallography over the past 6 decades. Struct Dyn 4:032102
75. Cullity BD (1978) Elements of X-ray diffraction. Addison Wesley, Boston
76. David WIF, Shankland K, McCusker LB, Baerlocher C (eds) (2002) Structure determination from powder diffraction data. IUCr monographs on crystallography, vol 13. Oxford Science Publications, Oxford
77. Warren BE (1990) X-ray diffraction. Addison Wesley, Reading
78. Lightfoot P, Tremayne M, Harris DM, Bruce PG (1992) Determination of a molecular crystal structure by X-ray powder diffraction on a conventional laboratory instrument. J Chem Soc Chem Commun 1992:1012–1103
79. Masciocchi N, Moret M, Cairati P, Sironi A (1993) Solving the structure of simple organometallic structures solely from X-ray powder diffraction data: the case of polymeric $[\{Ru(CO)_4\}_n]$. J Chem Soc Dalton Trans 3:471–475
80. Stokes AR, Wilson AJC (1942) A method for calculating the integral breadths of Debye-Scherrer lines. Proc Phys Soc 38:313–322
81. Stokes AR, Wilson AJC (1944) The diffraction of X-rays by distorted-crystal aggregates. Proc Phys Soc A 56:174–181
82. Warren BE (1969) X-ray diffraction. Addison Wesley, Reading

83. Warren BE, Averbach BL (1952) The separation of cold-work distortion on X-ray patterns. J Appl Phys 21:595–599
84. Lovesey SW (1984) Theory of neutron diffraction from condensed matter. Volume 1: Neutron scattering. Clarendon Press, Oxford University Press, Oxford
85. Piccoli PMB, Koetzle TF, Schultz AJ (2007) Single crystal neutron diffraction for the inorganic chemist – a practical guide. Comments Inorg Chem 28:303
86. Bau R, Drabnis MH (1997) Structures of transition metal hydrides determined by neutron diffraction. Inorg Chim Acta 259:27–50
87. Shull C (1995) Early development of neutron scattering. Rev Mod Phys 67:753–757
88. Ibberson RM, David WIF (2002) Neutron powder diffraction structure determination from powder diffraction data, Chapter 5. Oxford University Press, Oxford
89. Jeffrey GA (1997) An introduction to hydrogen bonding. Oxford University Press, Oxford, pp 235–245
90. Davies M (1947) The physical aspects of the hydrogen bond. Annu Rep Prog Chem 43:5–29
91. Hunter L (1947) The hydrogen bond. Annu Rep Prog Chem 43:141–154
92. Flierler U, Stalke D (2012) More than just distances from electron density studies. Struct Bond 146:1–20
93. Madsen AO (2012) Modelling and analysis of hydrogen atoms. Struct Bond 146:21–52
94. Engles B, Schmidt TC, Gati C, Schirmeister T, Fink RF (2012) Challenging problems in charge density: polar bonds and the influence of the environment determination. Struct Bond 147:47–98
95. Coppens P (2005) Charge density comes of age. Angew Chem Int Ed 44:6800–6811
96. Coppens P (1997) X-ray charge densities and chemical bonding. Oxford University Press, Oxford
97. Bader RF (1990) Atoms in molecules – a quantum theory. Oxford University Press, Oxford
98. Popelier PA (2014) The quantum theory of atoms in molecules. In: Frenking G, Shaik S (eds) The nature of the chemical bond revisited. Weinheim, Wiley, pp 271–308
99. Izyumov Yu A, Naish VE, Ozerov RP (1991) Neutron diffraction of magnetic materials, translated from Russian by Joachim Buchner. New York Consultants Bureau, New York
100. Seip HM (1973) Molecular structure by diffraction methods. Specialist Periodical Report, vol 1, p 1
101. Rankin DWH, Mitzel N, Morrison C (2013) Structural methods in molecular inorganic chemistry. Wiley, Weinhem
102. Hargittai I, Hargittai M (eds) (1988) Stereochemical applications of gas phase electron diffraction. VCH Publishers, Weinheim
103. Hilderbrandt RL, Bonham RL (1971) Structure determination by gas electron diffraction. Annu Rev Phys Chem 22(1):279–312
104. Rankin DWH, Robertson HE (2009) Davidson G, Ebsworth EAV (eds) Spectroscopic properties of inorganic and organometallic compounds, vol 40. Royal Society of Chemistry, London, pp 1927-1935
105. Wann DA, Rankin DWH, McCaffrey D, Martin JML, Mawhorter RJ (2014) Equilibrium gas-phase structures of sodium fluoride, bromide, and iodide monomers and dimers. J Phys Chem A 118:1925–1937
106. Wann DA, Reilly AM, Rataboul F, Lickiss PD, Rankin DWH (2009) The gas-phase structure of the hexasilsesquioxane $Si_6O_9(OSiMe_3)_6$. Zeit Naturforsch B 64:1269–1275
107. Hnyk D, Rankin DWH (2009) Stereochemistry of free boranes and heteroboranes from electron scattering and model chemistries. J Chem Soc Dalton Trans 4:585–599
108. Haaland A (2007) Molecules and models – the molecular structures of main group molecules. Oxford University Press, Oxford
109. Hargattai M (2000) Structures of metal halides in the gas phase. Chem Rev 100:2233
110. Haaland A (1988) Organometallic compounds of main group elements in stereochemical. In: Harattai I, Hargattai M, van Nostrand E (eds) Applications of gas phase electron diffraction, part B, Structural information for selected classes of compound. Wiley, New York

111. Haaland A (1989) Covalent vs dative bonds to main group metals – a useful distinction. Angew Chem Int Ed 28:992–1007
112. Haaland A (2016) Lewis and Kossel's legacy: structure and bonding in main group compounds in the chemical bond II. In: Mingos DMP (ed) Structure and bonding, vol 170, pp 1–70
113. Rankin DWH, Robertson HE, Danopoulos AA, Lyne PD, Mingos DMP, Wilkinson G (1994) Molecular structure of tetrakis(t-butylimido)osmium(VIII) determined in the gas phase by electron diffraction. J Chem Soc Dalton Trans 4:1563–1569
114. Thompson HB, Bartell LS, Bonham RA (1968) Inorg Chem 7:488–491
115. Hansen HW, Bartell LS (1965) Electron diffraction study of PF_5. Inorg Chem 4:1775–1779
116. Bartell LS (1972) Galloway G (ed) Molecular geometry in collected readings in inorganic chemistry, vol 2. Chemical Education Publishing, Easton, p 220
117. Gruene T, Wennmacher JT, Zaubitzer C, Holstein JJ, Heidler J, Fecteau-Lefebvre A, De Carlo S, Müller E, Goldie KN, Regeni I, Li T, Santiso-Quinones G, Steinfeld G, Handschin S, van Genderen E, van Bokhoven JA, Clever GH, Pantelic R (2018) Rapid structure determination of microcrystalline molecular compounds using electron diffraction. Angew Chem Int Ed 57:16313–16317

Struct Bond (2020) 185: 43–68
https://doi.org/10.1007/430_2020_76

Published online: 24 October 2020

Recent Developments in the Refinement and Analysis of Crystal Structures

Richard I. Cooper

Contents

Abstract Crystal structure refinement and analysis is a powerful method for determination of crystal structures and finds widespread application in determination of structures of crystals of small molecules and frameworks at atomic resolution. The independent atom model is used to describe atomic scattering for routine use, while more accurate aspherical scattering factors are increasingly available. The structure factor is presented as the Fourier transform of convolutions of scattering and probability densities in the crystal structure to clarify how aspherical scattering factors and alternative displacement probabilities can be introduced into refinement methods. Non-linear least squares fitting of the crystal structure parameters in the structure factor equations is described using matrix algebra notation which enables simple derivation of the extensions required for discussion of crystallographic restraints and leverage analysis. Finally, combined analysis of multiple single-crystal

R. I. Cooper (✉)
Chemical Crystallography, Department of Chemistry, University of Oxford, Oxford, UK
e-mail: richard.cooper@chem.ox.ac.uk

experiments is discussed highlighting the potential of refinement tools to extract useful information from joint X-ray and neutron data and from mixed ground-state and excited-state X-ray data from pump-probe experiments.

Keywords Atomic displacement parameters · Crystal structure refinement · Least squares · Leverage analysis · Restraints · Structure factor · Validation

Abbreviations

$\otimes$	Convolution operator
ADP	Anisotropic displacement parameter
DFT	Density functional theory
F_c^2	Calculated structure factor squared
FFT	FAST Fourier transform
F_o^2	Observed structure factor squared
FT	Fourier transform
IAM	Independent atom model
r.m.s.	Root mean square
SCF	Self-consistent field
TAAM	Transferable aspherical atom model
TLS	Translation-libration-screw
u(x)	Estimated uncertainty of a model parameter, x
λ	Wavelength
$\sigma^2(F_o^2)$	Estimated variance of the observed structure factor

1 Introduction and Background of Crystal Structure Refinement and Analysis

X-ray crystal structures have been determined from a diverse range of materials, spanning metals and minerals, through small-molecule organic and metal-organic compounds, including covalent and metal-organic extended frameworks, to crystallized proteins and even virus particles. This chapter focusses specifically on the determination of chemical information based on X-ray or neutron scattering techniques.

Structure solution algorithms based on direct methods or charge flipping can be run in parallel, in order to provide a high chance of finding a reasonable *trial structure solution*. Such solutions, despite being derived directly from the experimental data, should be considered as untested hypotheses. The accuracy of the structure can be improved, and the hypothesis tested, by fitting the model to the experimental data using *weighted non-linear least squares* algorithms. This approach provides several internal measures of quality including the goodness of fit to the data, estimates of precision of the parameters and convergence of the model.

1.1 Refinement and Analysis

For routine structure determinations, the independent atom model (IAM) is sufficient to explain X-ray scattering with enough accuracy that atom positions and displacements can be confidently determined; this structure analysis paradigm has not changed significantly since the early days of crystallography, nor have the underlying mathematical procedures. Nevertheless, increased computing power and optimized algorithms for linear algebra and Fourier transforms have increased the practical limits of the complexity that can be handled within a reasonable timescale.

Interesting developments in the field have occurred at the frontiers of chemistry and with increases in computational power, allowing excursions into new methods and applications of crystallographic modelling, many of which have been incorporated into routine analyses. The use of crystal structure refinement to handle difficult chemical and physical problems frequently leads to situations where extra information or assumptions are required to interpret or justify the results obtained. In extreme cases the structure refinement fit can no longer *independently* confirm that the structural model is correct, i.e. that it is a satisfactory approximation of the crystal structure itself. These conditions can range from lack of good-quality data due to data-limiting experimental conditions (e.g. such as high-pressure or high-temperature experiments), or intrinsic limitations of crystal quality (e.g. disorder, solvent voids and stacking faults), to hard-to-model crystal structures which contain multiple units related by non-crystallographic symmetry or modulation functions.

A key assumption for fitting a model by least squares minimization is that the errors are random and normally distributed about their mean values. If scattered X-rays are measured and processed carefully, the data will generally obey this criterion; however the model itself can violate it by introducing systematic errors into the fit.

1.2 Practical Elements of Refinement: The IAM Model

Using the independent atom model (IAM) to fit X-ray diffraction measurements requires an atomic scattering factor for each element; for very precise work, it may be necessary to use the scattering factor for the corresponding ion [1]. Scattering factors are computed from the spherically averaged electron density arising from relativistic Hartree-Fock atomic wave function calculations, or the analytical solution of the Schrödinger equation in the case of hydrogen atoms. The probability density for an atom in real space and its Fourier transform, the scattering power in reciprocal space, is shown in Fig. 1. The high densities close to the nucleus of the atom account for the majority of the scattering interactions, and small deviations away from the assumption of a spherically averaged X-ray scatterer, such as bonding or lone pair electron density, are ignored within the IAM approximation.

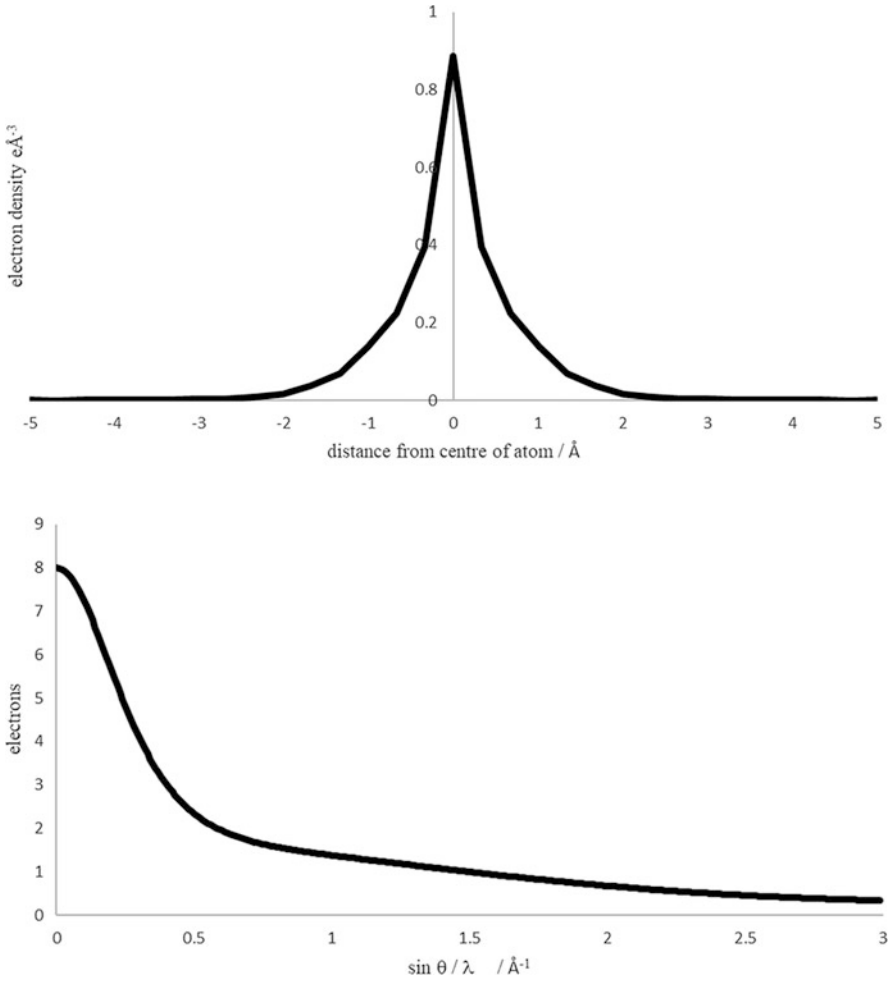


Fig. 1 Atomic electron density in real and reciprocal space. Left, the electron density as a function of distance from an atom; right, its Fourier transform – the scattering power in reciprocal space

Deviations from spherical atom density occur outside the regions of high electron density near the nucleus and correspond to more localized regions of electron density associated with chemical phenomena such as lone pairs and bonds. In reciprocal space these deviations from spherical symmetry will have the largest influence at low scattering angles, so a systematic bias is expected in low-angle X-ray measurements – those with low *hkl* indices. These small errors can manifest themselves as deviations in model parameters from their true values. The magnitude of the deviation is generally worse when high-angle X-ray data are not available [2].

The commonly used least squares weighting scheme (Eq. 1), e.g. as used in SHELXL [3], was designed to account for some systematic errors in the fit of the

model to data such as constant background [4]. The statistical least squares weights of $1/\sigma^2(F_o^2)$ are inversely proportional to the estimated variance of each observation. However, by inflating the variance using coefficients of functions of the observation magnitude, the weights can be systematically reduced for stronger observations. Strong X-ray observations are likely to be found at low angle, so this form of weighting scheme reduces some of the bias in the fit by increasing the relative importance of high-angle data, thereby accounting for shortcomings in the model itself:

$$w = \frac{1}{\sigma^2\left(F_o^2\right) + (aP)^2 + bP}; \text{where } P = \frac{2}{3}F_c^2 + \frac{1}{3}\begin{cases} 0, F_o^2 < 0 \\ F_o^2, F_o^2 \geq 0 \end{cases} \quad (1)$$

More advanced models of atomic scattering give more accurate estimates of structural parameters. These include (a) fitting multipole parameters directly to high-resolution data to describe the distribution of electron density; (b) using libraries of atomic charge density from theoretical calculations or high-resolution X-ray experiments, which can be included in a transferable aspherical atom model (TAAM) [5]; and (c) partitioning the computed electron density from density functional calculations on a molecule into aspherical atomic contributions, which are Fourier transformed to give the aspherical scattering factors for each atom. Following fitting of the structural parameters to X-ray data using (c), the process can be repeated to convergence, e.g. as implemented in Olex2 and Tonto [6]. These models all come at a cost in terms of convenience or experimental requirements: multipole refinement requires very high-resolution diffraction and data which is as free as possible from systematic errors; libraries tend to have regions of low coverage of some atomic species and chemical environments; and density functional methods are limited by speed and self-consistent field (SCF) convergence for heavier elements.

The details of obtaining atomic density from quantum chemical calculations, or from a database of transferable aspherical atomic models, are discussed in other chapters. A short discussion of the integration of these and other features into a standard least squares refinement is given below.

The Fourier transform of a crystal structure model is the predicted diffraction pattern of the crystal, subject to corrections for experimental geometry. The crystal structure model used to predict X-ray scattering describes the scattering density of each atom, combined with a coordinate within a crystal structure, and a function expressing its mean displacement from this position. The standard structure factor equation makes use of two key relationships:

1. The Fourier transform is a linear transformation: the *Fourier transform of the sum* of the scattering density of all atoms in a crystal is equivalent to *the sum of the Fourier transform* of each individual atom's scattering density:

Fig. 2 The convolution of atom position, atomic scattering density and displacement term give the overall atom scattering density

$$FT\left(\sum_j \rho_j\right) = \sum_j FT(\rho)_j \tag{2}$$

for j atoms, each with scattering density, ρ_j, centred at an atom position within a crystal. This allows the overall Fourier transform of a crystal to be expressed as a sum of the Fourier transforms of individual atoms, so that the model can be defined and optimized in a simple equation.

2. The convolution theorem demonstrates the equivalence of the *Fourier transform of a convolution of two or more functions* and the *product of the individual Fourier transforms of the functions* (Eq. 3):

$$FT(A \otimes B) = FT(A) \times FT(B) \tag{3}$$

A convolution operation, denoted ⊗, is defined as the integral, evaluated for all shifts, of the product of two functions, after one is reversed and shifted. A crystal structure model can be broken down into three key functions from which we can reconstruct the entire crystal scattering density using convolutions: firstly, a *position* relative to a periodic lattice, the position is represented by an infinitely sharp delta function; secondly, the *scattering density* of an atom at rest (for electrons a sharply peaked distribution and for neutrons an (effectively) infinitely sharp delta function with the scattering density of the relevant atomic nucleus; and, thirdly, a model for *displacement* of an atom from its mean position. The third term usually takes the form of a univariate or trivariate Gaussian distribution representing isotropic or anisotropic displacements. The convolution of a position, the atomic scattering density and the harmonic displacement function give the real space scattering density of an atom in a crystal (Fig. 2).

Using both of the above relationships, the structure factor (the Fourier transform of one unit cell) can be written as a sum of products of Fourier transforms of each of the individual functions above:

$$F_{hkl} = \sum_j f_j \times \exp\left[-8\pi^2 U_j \left(\frac{\sin\theta}{\lambda}\right)^2\right] \times \exp 2\pi i\left[\left(hx_j + ky_j + lz_j\right)\right] \tag{3}$$

where f_j, the atomic scattering factor, is the Fourier transform of the electron density of the j th atom; $\exp[-8\pi^2 U_j(\sin\theta/\lambda)^2]$, the isotropic displacement, is the Fourier

transform of a Gaussian distribution which expresses the time- and space-dependent harmonic displacements of an atom from its mean position; and, finally, $\exp 2\pi i [(hx_j + ky_j + lz_j)]$ is the Fourier transform of a delta function at the position of atom j.

The structure factor expression conveniently separates all of the contributions which makes replacing parts or extending the calculation quite straightforward. In order to replace f_j to produce models for fitting aspherical scattering density, as described above, only one multiplicative term of the expression needs to be changed. Note that any part of the structure factor expression that is not symmetrical about the origin, such as aspherical electron density, results in a complex number, and the overall structure factor computed for every observation is the product of these complex contributions.

Additional terms can be multiplied onto the existing structure factor equation to model scattering due to non-rectilinear motion of atoms. This corresponds to an additional convolution operation, for example, a function describing an arc segment can elegantly extend the existing atomic model to account for a nonatomic, non-rectilinear distribution. Examples of this approach include convolutions of isotropically distributed atomic density with a line, spherical shell or ring function allowing refinement of continuously disordered atom models [7]; the *hindered rotor* which convolves a 'normal' isotropic IAM atom with a function that distributes it around a ring with an adjustable hindering potential allowing modelling of scattering from groups with strong in-plane libration, e.g. coordinated cyclopentadienyl rings, benzene and CF_3 groups [8]; and directly derived skewed anisotropic and curvilinear distributions [9]. The first two cases are implemented in CRYSTALS [10] and have proved useful for modelling disordered solvent molecules or substituent groups, including spherical rotational disorder of tetramethylammonium ions in a series of tetrachlorometallate salts [11], the disordered equatorial density of spinning PF_6^- anions [12] and disordered guest solvent in a molecular cavitand [13].

1.3 Linear Algebra Description

Least squares optimization is an appropriate method to find the best fit of a crystallographic model to a set of measured data. The fit is computed between observed and computed structure factor magnitude squared, or sometimes observed and computed structure factor magnitude. Assuming suitable weighting of observations, the same best fit is calculated from both approaches.

The least squares method and its matrix algebra notation is described here and is useful to describe subsequent concepts.

The aim of a least squares refinement is to find the minimum value of χ^2, the sum of the squares of the differences between diffraction measurements under conditions $\boldsymbol{h_i}$ and their values predicted by the current model:

$$\chi^2 = \sum_i \left[y_i - f\left(\boldsymbol{h_i}, x_1, \ldots, x_j\right)\right]^2 \tag{4}$$

y_i is the value of the ith observation with index $\boldsymbol{h_i}$, and $f(\boldsymbol{h_i}, x_1, \ldots, x_n)$ is the prediction of that observation using Eq. (3) with model parameters x_1–x_n. These parameters are the atomic coordinates, displacements and any other refined parameters of a structural model.

The first derivative of χ^2 with respect to every parameter will be zero at the minimum, and the values of x that correspond to the best fit satisfy the set of equations:

$$\frac{d\chi^2}{dx_1} = 0, \ldots, \frac{d\chi^2}{dx_n} = 0 \tag{5}$$

For many problems, including fitting a crystallographic model to diffraction data, one or more of the derivatives of the χ^2 function contain non-linear functions of the other parameters in the set x_1–x_n, and as a result, the problem must be *linearized* and solved iteratively. Each structure factor equation can be linearized for each parameter using a first-order Taylor expansion, analytically replacing a complicated function with multiple exponential and trigonometric functions with a simple straight-line equation which follows the tangent of the χ^2 function at the current value of each parameter.

Unlike linear problems, an approximate model is required as the starting point, since a value is required for each parameter to carry out the Taylor expansion. The starting model can be generated from any structure solution method, from analogous crystal structures, or from pure inspiration – the least squares process, and subsequent validation, is the test of its correctness.

A further consequence of the approximation made in linearization is that the adjustments made to the parameters will move towards the minimum of Eq. (4), but multiple iterations will be required to ensure that they have converged to the best solution.

Matrix algebra notation represents these equations concisely and can be extended easily to demonstrate incorporation of restraints and other analyses.

The vector of shifts to be applied to each of the parameters is denoted $\mathbf{\Delta X}$ and contains one element for each least squares parameter. $\mathbf{\Delta Y}$ is a vector containing one element for the difference between every observation and its calculated value, sometimes known as *the residual*. Each row of the matrix $\mathbf{A}$ contains the derivative $\frac{d\chi^2}{dx_j}$ for one observation with respect to each parameter in turn. The set of observational equations for the linearized problem may be written as

$$\mathbf{\Delta Y} = \mathbf{A}\ \mathbf{\Delta X} \tag{6}$$

Setting the derivatives of the χ^2 function to zero and solving these equations with appropriate weights yields the *normal equations* which give the shifts, $\mathbf{\Delta X}$, to parameters which will step the parameters closer to the nearest minimum in χ^2:

$$\left[\mathbf{A}^{\mathbf{T}}\mathbf{W}\mathbf{A}\right]^{-1}\mathbf{A}^{\mathbf{T}}\,\mathbf{W}\,\boldsymbol{\Delta}\mathbf{Y} = \boldsymbol{\Delta}\mathbf{X} \tag{7}$$

W is a diagonal matrix containing a weight for each observation. The quantities in the matrices on the left side are all known values, and the required shifts may be computed using straightforward matrix multiplications and inversion.

1.4 Computing Power

For the three or four decades up to 2005, computer clock speeds increased exponentially each year from sub-Mhz until they reached approximately 2 GHz. This trend supported by equivalent increase in power and capacity of other components meant that code would just run faster each time it was ported to or installed on new hardware. However, since 2005, clock speeds have stalled, and improvements have been delivered by smaller transistors (Moore's law) and multi-core processors. To take advantage of these developments, software code often requires significant reorganization to allow parts of substantial calculations to run in parallel.

High-performance mathematical libraries, in particular those based on BLAS and LAPACK, make use of standardized cross-platform interfaces for linear algebra operations which allows developers to focus on mathematical problems rather than the computer science [14–16]. These libraries can outperform manual optimization of code by several orders of magnitude and tend to scale to larger problems much more efficiently. Most implementations also take advantage of multi-core processors when available, speeding up code and saving developers from further work organizing the parallelization of calculations. Figure 3 illustrates an example of the advantages of switching an existing least squares program to a high-performance library: for a problem with 1,261 least squares parameters, during the computation of

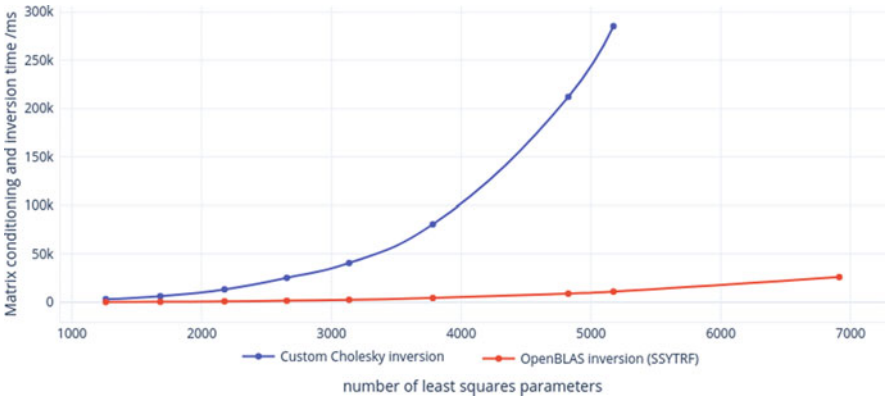


Fig. 3 Comparison of matrix inversion time as a function of least squares problem size. Plot generated data in reference [17]

$[\mathbf{A^T WA}]^{-1}$, a LAPACK library routine outperforms a hand-optimized Cholesky matrix inversion routine by a factor of 18. By the time the problem size reaches just over 5,000 parameters (approx. 500 non-hydrogen atoms), this factor has risen to 26, with wall-clock times of 5 min for a single least squares cycle using the original routine and just over 10 s for the library [17]. Rapid structure factor calculation and matrix inversion enables testing of many more hypotheses than would have been possible 20 years previously.

The Fourier transform is another commonly used operation in crystallographic analysis that has benefitted from abstraction and optimization of libraries to carry out the operation. In addition to being used for generating the scattering density for peak searching and visualization, Fourier transforms can be used to convert atomic scattering density into custom aspherical scattering factors and are an integral part of the SQUEEZE algorithm for correction of unmodelled scattering from solvent in voids [18].

As crystallographic utilities are developed and improved, off-the-shelf optimized and efficient implementations of the discrete Fourier transform such as FFTW3 [19] are outperforming and replacing manually optimized Fourier calculations.

2 Determination of Absolute Configuration During Crystal Structure Refinement

The Fourier transform of a three-dimensional non-centrosymmetric real valued function is centrosymmetric. The intensity of an X-ray diffraction pattern formed by elastically scattering X-ray photons from a material is proportional to the square of the Fourier transform of the electron density of the material (a real valued function). Therefore this pattern of intensity will also be centrosymmetric, a relationship known as *Friedel's law*. This fact is consistent with the fact that a structure factor, F(*hkl*), provides information about the distribution of electron density sampled by a plane wave of a particular direction and spacing, while its centrosymmetric equivalent, $F(\overline{hkl})$, provides information about the same scattering density relative to the same plane wave, just defined in the opposite direction.

The symmetry of a diffraction pattern from a non-centrosymmetric material which obeys Friedel's law is the *Laue symmetry* of the crystal. This is a supergroup of the *point symmetry* of the crystal, formed by the addition of a centre of inversion.

Resonant X-ray scattering from one or more atoms can break the symmetry between the structure factor magnitudes of a pair of reflections related by inversion. This effect was observed experimentally in early X-ray experiments [20] and is now routinely used to determine the *absolute structure* of materials. The resonant scattering effect is generally stronger for heavier elements and longer X-ray wavelengths. It is included in the model by introducing a complex atomic scattering factor with a small imaginary component.

Table 1 Examples of indices of symmetry related reflections in monoclinic (b-unique) space group $P2_1$

hkl	$\bar{h}\bar{k}\bar{l}$	Inversion
hkl	$\bar{h}k\bar{l}$	Twofold rotation
hkl	$h\bar{k}l$	Inversion + twofold rotation
$\bar{h}k\bar{l}$	$h\bar{k}l$	Inversion

A distinction is sometimes made between structure factor intensities related by inversion, known as Friedel pairs, and intensities related by an inversion plus any nonidentity operation of the point group symmetry of the crystal, which are may be referred to as Bijvoet pairs. However, this distinction is not universal, and the terms are used interchangeably in the literature. Table 1 lists four pairs of general structure factor indices for a structure in space group $P2_1$. With the exception of the second line, all are related by an operation which includes inversion and is not part of the space group. The differences in these three cases can be used to help determine the absolute structure of a material.

In crystal structures containing no operations of the second kind (inversion, mirror, or glide), assignment of *absolute structure* allows the direct assignment of the chirality of the chemical structure itself. For the class of non-centrosymmetric materials that contain a mirror or glide plane, determining whether the direct structure, or its inverse, gives rise to a diffraction pattern will determine the direction of a polar axis within the structure, but this information is rarely likely to be of interest in a chemical crystallography application.

Methods for determining absolute structure from X-ray diffraction data have undergone several developments, which have improved usability and the robustness and precision of its assignment. Early approaches directly compared the crystallographic R-factor of a structure fit, against the R-factor of its inverse, either directly or using the Hamilton R-factor ratio test [21]. The method works reliably when there is significant resonant scattering signal, but the approach could give inconsistent results for weak resonant scattering signals or in the presence of uncorrected systematic errors. The Rogers η parameter directly refined a coefficient of the imaginary component of the resonant scattering factor [22], ideally converging to either 1 or −1, indicating the correct or inverted absolute structure of the refined model, respectively.

Although refined as a continuous variable, the η parameter has no physical meaning when its value is not exactly ±1. Flack showed that the η parameter could potentially have an unexpected double minima resulting in an instability if refinement is started at 0. As an alternative he proposed direct competitive refinement of the volume fraction of a structural model, $1 - x$, and the volume fraction of its inverse, x, (Eq. 8) resulting in the ability to characterize crystals containing domains related by inversion and at the same time obtain a reliable estimate of the uncertainty [23]. The parameter x is referred to as the *Flack parameter*, and its standard uncertainty is denoted $u(x)$:

$$|F_{\text{twin}}(hkl)|^2 = (1-x)\left|F_{\text{single}}(hkl)\right|^2 - x\left|F_{\text{single}}\left(\overline{hkl}\right)\right|^2 \quad (8)$$

where $|F_{\text{single}}(hkl)|^2$ is the structure factor magnitude computed from the structural model without twinning by inversion.

The magnitude of $u(x)$ must be sufficiently small to enable confident interpretation of x. Values of $u(x) < 0.1$ have been shown to indicate sufficient precision to trust the value of x if the enantiopurity (but not the chirality) of the material is already known, and values of $u(x) < 0.04$ are required to have confidence in the value of x when enantiopurity cannot be verified [24]. If $u(x)$ is sufficiently small, a value of x of close to 0 indicates that the structural model has the same absolute structure as the crystal, and a value of close to 1 indicates the structural model is the inverse of the crystal. Values significantly different ($>2u(x)$) from 0 and 1 indicate a crystal which is twinned by inversion.

Research has been directed at improving the estimates of uncertainty associated with the Flack parameter, by formulating expressions for the Flack parameter that are based directly on observed (D_{obs}) and calculated (D_{single}) differences between Friedel pairs of acentric reflections:

$$D_{obs}(hkl) = |F_{obs}(hkl)|^2 - \left|F_{obs}\left(\overline{hkl}\right)\right|^2$$

$$D_{\text{single}}(hkl) = \left|F_{\text{single}}(hkl)\right|^2 - \left|F_{\text{single}}\left(\overline{hkl}\right)\right|^2$$

The slope of the straight-line fit of D_{obs} against D_{single} is $1 - 2x$, where x is the original Flack parameter. Values of $u(x)$ determined by this method were shown to be on average 3–4 times smaller (i.e. more precise) than those obtained by traditional refinement of x in full-matrix least squares [25].

A related approach uses a quotient formula which removes a dependence on the crystallographic scale factor and gives results that are more or less equivalent [26]:

$$Q_{obs}(hkl) = \frac{|F_{obs}(hkl)|^2 - \left|F_{obs}\left(\overline{hkl}\right)\right|^2}{|F_{obs}(hkl)|^2 + \left|F_{obs}\left(\overline{hkl}\right)\right|^2}$$

$$Q_{\text{single}}(hkl) = \frac{\left|F_{\text{single}}(hkl)\right|^2 - \left|F_{\text{single}}\left(\overline{hkl}\right)\right|^2}{\left|F_{\text{single}}(hkl)\right|^2 + \left|F_{\text{single}}\left(\overline{hkl}\right)\right|^2}$$

The straight-line fit of Q_{obs} against Q_{single} is $1 - 2x$.

2.1 Highly Disordered Resonant Scatterers

A difficult case for absolute structure determination occurs when a significant proportion of the resonant scattering atoms are very disordered and cannot be

included in the atomic model of the crystal structure, for example, when solvent molecules occupy a void within a crystal structure with no strong interactions to order them relative to the rest of the structure. Extreme cases are light atom molecular structures which contribute very little resonant scattering, but which contain very disordered solvent with some strong resonantly scattering atoms, such as the chlorine atoms in dichloromethane.

A popular treatment of disordered solvent regions is the SQUEEZE procedure implemented in PLATON [18]. SQUEEZE computes a contribution to the calculated scattering from a crystal structure model via a discrete Fourier transform of the residual electron density from the disordered region. The process is iterative so that the residual density estimate can benefit from the improved model. However, the corrected model cannot normally account for resonant scattering from atoms in the disordered region as the computed electron density has no atomic resonant scattering. Straightforward determination of absolute structure from crystals where the strongly resonantly scattering atoms are not resolved has therefore not been possible.

Application of a resonant scattering correction to the SQUEEZE contributions from the disordered regions, weighted by a function of resonant scattering terms of the missing atoms, has allowed recovery of absolute structure information using both conventional Flack x refinement and other post-refinement methods [27].

It should be noted that resonant scattering effects are not observed for most atom types in neutron diffraction; however the dynamical refinement methods used to model multiple scattering in electron diffraction analysis can determine the absolute structure of light atom molecules [28].

3 Embedding Information Using Crystallographic Restraints

Crystallographic restraints are a convenient means to introduce additional information (*subsidiary conditions*) when fitting crystal structure model parameters to diffraction data using least squares [29].

For conventional crystallographic least squares, the function to be minimized is the sum of the squares of the residuals, $\chi^2 = \sum_i w_i[y_i - f(\mathbf{h_i}, \mathbf{x})]^2$, where y_i is the observed value of the i-th data point, measured under conditions $\mathbf{h_i}$, and $f(\mathbf{h_i}, \mathbf{x})$ is the predicted value of the same observation given a set of model parameters, $\mathbf{x}$. Each residual is assigned a suitable weight, w_i, computed from an estimate of the standard deviation of the observation. A supplementary sum of residuals allows parameter restraints to be included as additional weighted equations which relate simple functions of the model parameters to the expected values of those functions:

$$\chi^2 = \sum_i w_i [y_i - f(\mathbf{h_i}, \mathbf{x})]^2 + \sum_j w_j \left[y_j - f_j(\mathbf{x})\right]^2 \tag{9}$$

where y_j is the expected value of a function of the model parameters $f_j(\mathbf{x})$ (e.g. a bond distance or the deviation from the average of a set of bond distances) and w_j is computed as the reciprocal of the square of the estimated standard deviation of y_j.

Using the matrix algebra notation introduced earlier, the weighted observational equations can be augmented with additional restraints as follows and solved for $\mathbf{\Delta X}$ using the same operations:

$$\begin{pmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{\Delta Y} \\ \mathbf{\Delta R} \end{pmatrix} = \begin{pmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{A} \\ \mathbf{S} \end{pmatrix} \mathbf{\Delta X} \tag{10}$$

where $\mathbf{\Delta R}$ is a vector of the differences for each restraint target and its current value in the parameters and the rows of $\mathbf{S}$ contain the derivatives for each restraint with respect to each parameter.

The purpose of introducing information differs depending on the stage of model development: (i) stabilizing refinement during the early stages of model building and testing, or (ii) providing necessary additional information that is not present in the scattering data. The first type compensates for a poor or incomplete model, while the latter type is likely to be required in the case of noisy or low-resolution scattering data, or in cases of pseudo-symmetry, where subsets of the crystallographic parameters are determined by weak and noisy subsets of the scattering data.

It is useful to be aware that restraints are treated as additional observations on a par with the experimental measurements in a weighted non-linear least squares fit, and therefore each must be supplied with an appropriate weight, which reflects confidence in the restraint target value, relative to the weights of the hundreds or thousands of scattering observations, which are usually supplied by the data reduction software used in the experiment.

Most crystallographic refinement software provides a common set of restraint functions which can be used to add relevant information about the relationships between model parameters, e.g. a *distance* restraint which relates the six coordinate parameters of two atoms. Other examples of geometric restraints include the *valence angles, torsion angles, planarity and chiral volume.* Each of these restraints can be equated either with an *absolute value* which defines the ideal value that the parameters should adopt or a *relative value*, which will link one restraint function value to another, e.g. restraining three distances to be equal could be used to tidy the C-F bond lengths in a CF_3 group. These relative restraints have the advantage that they do not impose possibly arbitrary values on geometry, but merely relate data about one geometric feature to another under some fairly straightforward chemical assumptions.

3.1 Displacement Parameter Restraints

The mean square *displacement* of an atom from its mean position is a commonly determined quantity in small-molecule single-crystal refinements and will sometimes be a single parameter defining the mean square displacement of an atom from its position (isotropic displacement), or a 3x3 covariance matrix (six independent parameters) defining the mean square displacements of an atom in three dimensions (anisotropic displacement).

Isotropic displacement parameters, where used, can be restrained to be equivalent to neighbouring isotropic displacement parameters, or to a 'U_{equiv}' value when the neighbouring atom has an anisotropic set of displacement parameters. This type of restraint is most commonly used to provide sensible values for the isotropic displacement parameters of a hydrogen atom attached to an organic ligand or molecule. A multiplier of 1.2–1.3 may be used to reflect the expectation that a H atom will have a larger displacement than a heavier atom to which it is attached due to the difference in mass.

Anisotropic displacement parameters (ADPs) often constitute approximately two-thirds of the total number of refined parameters in a structure analysis, yet compared to the number of geometrical restraints available, there are relatively few displacement parameter restraints available. Of those in routine use, the Hirshfeld rigid-bond restraint [30] is the most physically valid and useful. It is based on the Hirshfeld criterion that the components of displacement parameters along the direction of a bond connecting two atoms should be approximately equal. It is also routinely applied to the direction between 1,3 connected atoms due to the relative rigidity of bond angles in molecules.

A further relationship commonly imposed upon displacement parameters of connected or overlapping disordered atoms is to set corresponding parameters of the anisotropic displacement matrix to be equal. This is also often applied as a constraint, but in both cases has the effect of making two atoms have the same or approximately the same mean square displacements from their average positions. This type of relationship is not physically meaningful and is rarely empirically observed between bonded atoms, so it is used sparingly, though it is occasionally useful for relating displacements of partially occupied atoms that are located on the same site or very close together in the unit cell representation of the structure.

A significant development in displacement parameter restraints was the introduction of the RIGU restraint, implemented in SHELXL [31]. This introduces the additional well-established empirical relationship between the covariances of displacement parameters perpendicular to the vector between two bonded atoms (Fig. 4b).

Both the Hirshfeld rigid-bond restraint and the RIGU enhancement make use of a restraint which can be readily expressed in a Cartesian coordinate system, which is conveniently aligned with the local geometry – in this case the *Z* axis is aligned with the interatomic vector. To write general expressions for restraints using similar coordinate systems, a transformation operation **T** is defined which transforms from

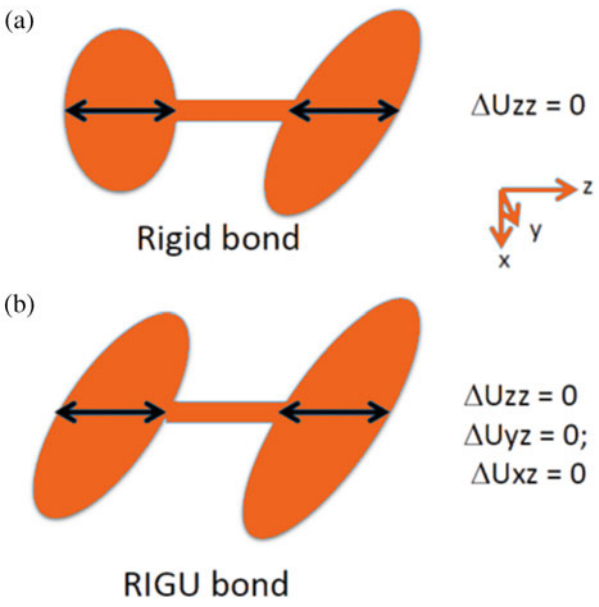


Fig. 4 Relationships between anisotropic displacement parameters in a local orthogonal coordinate system where the *z*-direction is aligned with the interatomic vector (**a**) rigid-bond restraint and (**b**) RIGU restraint. Subscript *zz* denotes quantities related to variance in the z-direction, and subscript *xz* denotes quantities related to the covariance between the *x*- and *z*- directions

the anisotropic variance-covariance matrix **U**, which is aligned with crystallographic reciprocal axes, but has units of Å^2, into a local Cartesian basis L aligned with local geometry:

$$[\mathbf{U}]_{\mathbf{L}} = \mathbf{TUT}^{\boldsymbol{T}}$$

The operation T can be broken down into three sequential operations, $\mathbf{T} = \mathbf{RAN}$. The first is a scaling matrix **N**, with reciprocal cell lengths on the diagonal – this converts the crystallographic U matrix into a dimensionless matrix. Second is an orthogonalization matrix **A**, which transforms the basis into an arbitrary but well-defined Cartesian coordinate system, and finally a pure rotation matrix **R** which relates the arbitrary Cartesian system to a local Cartesian system of interest.

Restraints can be defined which relate the components of $[\mathbf{U}]_{\mathbf{L}}$. For example, to force one of the principal axes of the distribution to align with the *X* direction in the local coordinate system, there should be no covariance between the displacements along the *X* direction and either of *Y* or *Z* directions; thus, $[\mathbf{U}]_{\mathbf{12,L}} = 0$ and $[\mathbf{U}]_{\mathbf{13,L}} = 0$. The 1,2 and 1,3 elements of the derivative of the $[\mathbf{U}]_{\mathbf{L}}$ matrix with respect to each of the crystallographic Uij parameters can be obtained by considering the effect of the transformation **T** on the relevant elements of the **U**ij matrix. This will reveal a linear combination of the crystallographic parameters that will be restrained to zero. The relative contributions depend only on the transformation, **T**.

In this construction, the Hirshfeld rigid-bond restraint is defined by choosing a local basis with the *Z* axis aligned with the interatomic vector. *X* and *Y* are mutually perpendicular to each other and perpendicular to *Z*, but their choice is otherwise

arbitrary. The element $[\mathbf{U}]_{33,\mathbf{L}}$ of both atoms is restrained to the average value of the two. Similarly the RIGU restraint can be defined by using the same basis and also restraining both $[\mathbf{U}]_{13,\mathbf{L}}$ and $[\mathbf{U}]_{23,\mathbf{L}}$ values to be equal to their respective average for each atom.

The range of ADP restraints can then be extended to suit particular chemical problems for which the traditional ADP restraints are not adequate [32]. These include (1) restraining a displacement parameter to make two principal axes perpendicular to a bond or angle bisector, to make it consistent with a librational motion in a molecule; (2) restraining one principal axis be perpendicular to a plane, to ensure consistency with libration in that plane; and (3) restraining two displacement parameters to be equivalent, but each within its own local coordinate system defined by the local bond topology, giving a much more physically reasonable restraint than traditional restraining or constraining pairs of Uij parameters to be equal. Restraints can also be defined to make the product or mean of the principal axes of a set of atoms the same, which can offer some relatively unbiased control of poorly determined sets of displacement parameters.

Finally, and most generally, ADP restraints can be derived from a TLS analysis of four or more atoms that form a rigid group [33]. The TLS model fits 20 independent translation, libration and screw tensor parameters to the existing refined anisotropic displacement parameters of the atoms. The displacement parameters can then be restrained to give the best fit to the computed TLS model, which will tend to make them well behaved across the whole rigid group.

The Simple Hydrogen Anisotropic Displacement Estimator (SHADE) server [34] combines the rigid group TLS analysis with empirical information about internal molecular displacements of hydrogen atoms in order to produce realistic estimates of the anisotropic displacement parameters of the hydrogen atoms in molecules. This approach can be used to improve the fit of a model to diffraction data in cases where there is a significant deviation from isotropic displacement of hydrogen atoms, but there is insufficient data to refine the hydrogen displacement parameters directly, such as low scattering signal in X-ray diffraction experiments, or low data to parameter ratios in neutron diffraction experiments.

4 Validation of Structure Refinements

Validation of crystal structures has grown from a spare-time community effort during the twentieth century. Researchers would sift through results of crystallographic analyses to find incorrectly assigned space groups [35, 36], occasionally the identification of an unlikely structure based on previous chemistry [37] or suspicious intermolecular interactions [38]. More recently a fully automated prepublication service supported by accessible results *and* data [39] has been developed, which aims to guard against mistakes in an analysis. It has also identified some cases of deliberate fraud [40], but fortunately this has been very rare. In addition to checking for incorrect syntax or missing data, the checkCIF suite of tests enforce tried and

tested limits on some statistical descriptors of the least squares fit (R-factor, goodness-of-fit, scattering density statistics), as well as checking for reasonable chemical and physical consistency of the crystal structure model such as balanced charges, bond lengths and angles, intermolecular interactions, coordination of ions, hydrogen bond geometry and displacement parameter consistency.

While these checks provide a degree of confidence in the correctness of a crystal structure determination, they are no substitute for a thorough understanding of a crystal structure model and its fit to the data. Simple plots of the least squares residual can reveal systematic errors and outliers in the data, which may not be caught by a one-size-fits-all service such as checkCIF. In particular Henn and Meindl have demonstrated that common crystallographic statistics, including the goodness-of-fit measure, may obscure systematic errors in data and have proposed alternative measures that are more robust [41, 42] and fractal dimension plots of residual scattering density to test for systematic biases in the data or model [43]. These developments have been made in the context of analysis of charge density data sets, but can equally be applied to routine IAM models to detect uncorrected data or under-parameterized models.

Complementary to these statistical checks are some methods which are not as simple to automate, either due to a relatively high computational cost or because the results require some careful inspection and interpretation. Approaches of both types are outlined below: leverage analysis for testing the appropriate application of crystallographic restraints in refinement and application of DFT calculations to crystal structure validation.

4.1 Leverage Analysis for Validation

A crystal structure is described by a set of model parameters, the optimal values of which are found using least squares fitting to experimental diffraction measurements *and* empirical restraints. In addition, the method yields a variance-covariance matrix which, when appropriately scaled, gives an estimate of the variance of each parameter, usually reported as a standard uncertainty. This matrix can be used to derive estimates of standard uncertainties of functions of the model parameters, such as bond lengths, angles, etc., by taking into account the covariance between parameters.

The leverage of each observation is defined as the influence that it has – through the structural model – on its own fitted value. For example, if the magnitude of one observation, F_{obs}, in a structure refinement is doubled, the leverage tells us what will happen to the value of F_{calc}. At one extreme a leverage of zero indicates that the model cannot respond to the change in F_{obs} in a way that will improve the fit for that reflection. Such reflections are rare but provide a useful internal check on the quality of the model. At the other extreme, a value of 1 indicates that the F_{calc} will follow the change in F_{obs} exactly, meaning that no other observations have any influence on its fitted value.

Leverage values are obtained from the projection matrix **P**, which transforms the differences between observed and calculated observations, **ΔY**, to the differences caused by the current set of parameter shifts, **AΔX**:

$$\mathbf{P}\,\mathbf{\Delta Y} = \mathbf{A}\,\mathbf{\Delta X} \tag{11}$$

Substituting the expression for **ΔX** from the normal equations (Eq. 7) yields

$$\mathbf{A}\left[\mathbf{A}^{\mathbf{T}}\mathbf{W}\mathbf{A}\right]^{-\mathbf{1}}\mathbf{A}^{\mathbf{T}}\,\mathbf{W}\,\mathbf{\Delta Y} = \mathbf{A}\,\mathbf{\Delta X} \tag{12}$$

By inspection **P**, the projection matrix is $\mathbf{A}\,[\mathbf{A}^{\mathbf{T}}\mathbf{W}\mathbf{A}]^{-\mathbf{1}}\,\mathbf{A}^{\mathbf{T}}\,\mathbf{W}$ [44, 45]. The **P** matrix has a row and column for every observation, but the leverages are the leading diagonal terms P_{ii}, so the computation and storage can be carried out efficiently. Its calculation is based on the **A** and **W** matrices only – the observations or residual differences are not used – underlining the fact that leverage is a feature of the model, not of the values of the observations.

Leverage values of individual X-ray data points tend to have few useful applications: Prince extended his analysis above to compute the improvement to individual parameter standard uncertainties caused by remeasuring individual X-ray observations. Picking the most useful reflections to remeasure is not immediately compatible with the way modern diffractometers collect data, since they sweep through large regions of reciprocal space, collecting multiple parts of the reciprocal lattice at the same time.

However, leverage finds an application in assessing the appropriate use of restraints in a structure refinement. In Sect. 3, restraints were defined as additional observational equations with a user-supplied standard uncertainty, which is used to compute a weight for the observation in the least squares. As a result of this mathematical equivalence of X-ray data and restraints, the leverage analysis above can equally be applied to a restraint to test whether it has (a) no influence, (b) reasonable influence or (c) far too much influence, on the parameters which it is restraining.

The result of two refinements of the same structure is shown in Fig. 5, illustrating how the leverages of restraints change as other data are removed. In this case 30 rigid-bond restraints are applied to 1,2 and 1,3 distances (with standard uncertainties of 0.002 $Å^2$ and 0.005 $Å^2$, respectively). On the left is a 'normal' refinement – the restraints have a very low average leverage, and so we can conclude that the X-ray data determines the values of the restrained parameters quite satisfactorily. On the right, only low-angle X-ray data is included, and the restraints now almost completely determine the values of the displacement parameters with an average leverage of around 0.8.

Leverage is therefore a useful metric of the influence of each restraint and allows an analyst to check the source of information determining a parameter before drawing any conclusions about its value.

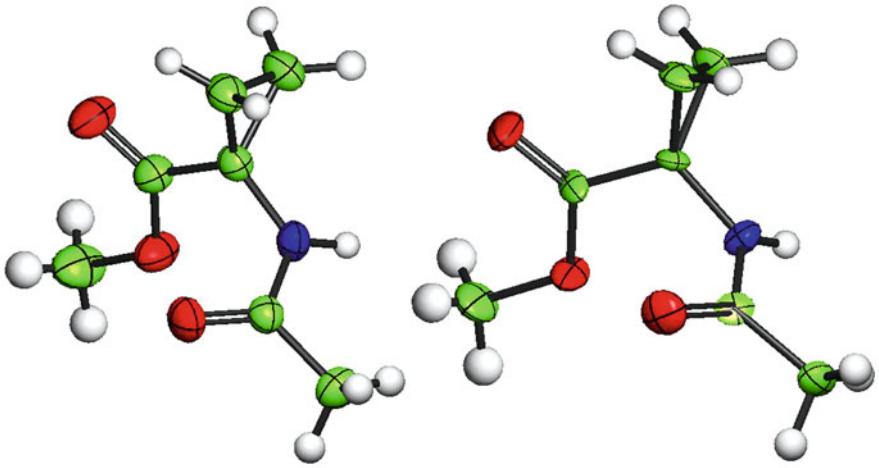


Fig. 5 Left, 729 X-ray data, 104 parameters, 30 rigid-bond restraints. R1 = 0.039 Rw = 0.113; restraint average leverage below 0.1 (max: 0.3); right, 136 X-ray data (low angle), 104 parameters, 30 rigid-bond restraints. R1 = 0.030 Rw = 0.077; restraint average leverage 0.8 (max, 0.95)

4.2 Validation with DFT

Periodic plane-wave density functional theory (DFT) is a powerful tool for computing properties of crystalline materials and finds applications alongside crystallographic investigations of materials. Common applications include surveying an energy minimum, or rationalizing geometry of structurally important atoms, for example, proton positions in short strong hydrogen bonds [46], and ranking of putative models in ab initio crystal structure prediction work [47]. DFT and phonon calculations can provide missing information, such as realistic anisotropic displacement parameters, in structural models fitted to low- or medium-resolution experimental data [48].

In principle DFT calculations can be used to validate molecular crystal structure results, either at the point of deposition or as a retrospective survey to provide confidence in historically reported structures. The latter application does not even require access to experimental diffraction data, so this could be applied to reports of structures going back over a century.

Geometry optimization of experimental crystal structures using dispersion-corrected DFT can give a measure of the agreement between experiment and theory. Application of this method to 241 organic crystal structures from a single volume of *Acta Cryst* E was carried out and revealed that the average r.m.s. Cartesian change in coordinates was 0.095 Å [49]. Based on analysis of outlying values, the authors suggest that a r.m.s. Cartesian displacement of atom positions of greater than 0.25 Å indicates either an error, such as an incorrect structure or incorrectly modelled disorder, or in one case an interesting temperature-dependent effect which resulted in a concerted shift of a whole molecule by 0.5 Å. Experimental verification

confirmed the existence of this shift between the 100 K and room-temperature structures, presumably governed by intermolecular interaction potentials.

There are limitations to automation of this approach including the following: heavier atoms increase the problems associated with SCF convergence, so it is generally applicable only to molecular organic crystal structures; unmodelled disorder often results in atom coordinates located between two alternative disorder sites, with large displacement parameters which correspond to a minimum in the fit to the crystallographic data, but do not correspond to an energy minimum for the system; and finally modelled disorder requires manual intervention to consider each conformation in turn in order to validate.

5 Horizons: Analysis of Multiple Experiments

Understanding properties of and characterizing new materials will often rely on a synthesis of data and observations from multiple experiments and theoretical calculations. In crystallographic analyses, use of restraints is one method for including external information in a least squares fit.

An area which is gradually developing is combining multiple sources of data into one model where multiple diffraction techniques have been employed – e.g. a combined X-ray and neutron model can take advantage of the accurate unit cell and heavy atom positions from X-ray diffraction data, while the neutron diffraction data can determine the hydrogen atom parameters, the contrast between scattering of elements with similar atomic number and magnetic ordering. This approach has been successfully applied to analysis of powder diffraction experiments [50–52], in part because instrument parameters are included in the refined model, making a multi-data set refinement a simpler extension of existing software. For single-crystal experiments, combined studies are possible using software packages such as Jana [53], TOPAS-Academic [54] and GSAS-II [55] and have been used to determine the distribution of neighbouring metal ions in high pressure cells, where neither technique alone provided sufficient information [56].

There is a tension in this combined approach, which hinders its routine application, because X-rays and neutrons probe different scattering density in the crystal structure. As well as having different magnitudes, the two densities do not necessarily have coincident extrema (hydrogen has a negative scattering length, so in this discussion, the extreme values have opposite signs: a maximum for X-ray scattering density and a minimum for neutron scattering density). For chemical crystallography, large differences of between about 0.09 Å and 0.16 Å occur between the positions of the extrema of scattering density of the hydrogen atom [2]. The extrema correspond to the mean of the distribution of the nuclear position in a neutron experiment, and the average of the electron positions in an X-ray experiment, which may be offset from the nuclear position when the atom is bonded to another.

Powder diffraction tools have long incorporated multi-model fits to an experiment, in order to account for multiple material phases in a powder sample. In the

single-crystal case, we require multi-model refinements which compute the scattering of, and are fitted to, different types or wavelengths of radiation.

Fitting independent models to separate data sets is equivalent to analysing the data sets in isolation; however crystallographic constraints and restraints can be used to define relationships and link common parameters between two or more models. For example, in a joint X-ray and neutron study, a common set of coordinates of non-hydrogen atoms can be used to model scattering for both sets of experimental data, while separate sets of hydrogen atoms can be used for each radiation type. Nevertheless, the hydrogen atom positions are not entirely independent: each pair of hydrogen atom positions can be constrained to lie on a common vector with the atom they are bonded to.

The approach can also be applied to modelling excited-state geometrical transformations in single crystals: a large part of the geometry of the structure is common to both ground state and excited state, and the diffraction data from the excited state, obtained by pump-probe methods, is often noisy and insufficient on its own to support full refinement of the excited-state structure.

The mathematical tools required to carry out these types of analyses are embedded within many currently available packages, but demonstration of their use and developments of protocols and tools for applying them, appropriate to the nature of the problem, are still required in order to bring them into mainstream use.

6 Summary

This chapter details some recent developments in the refinement and analysis of small-molecule crystal structures. The IAM model coupled with harmonic atom displacements is the cornerstone of small-molecule X-ray analysis and will undoubtedly remain important for standard structure determination and structures for a long time, especially for structures with atom environments for which aspherical scattering factors are not known and cannot easily be calculated. Section 1.2 shows how alternative scattering models and probability distributions can be incorporated into a routine refinement.

Section 1.4 highlights the potential for crystallographic studies to take advantage of improvements in computing power, not only for speeding up analyses but potentially to test alternative hypotheses for models, for example, in building and comparing multiple models of disordered chemical groups and molecules.

Recent developments in absolute structure determination from resonant scattering are outlined in Sect. 2, including details of the determination of the Flack parameter by analysis of Friedel pairs of X-ray reflections, instead of least squares fit with the other model parameters.

Sections 3–5 outline developments within the framework of traditional crystallographic constraints and restraints to provide more physically meaningful crystal structure models. Applications include restraining atomic displacement parameters and linking together models which are fitted against different sources of

experimental data. The validity of models can be assessed using traditional least squares metrics. Increasingly, DFT calculations can provide a theoretical check on results, while leverage analysis can check the influence of restraints on the fit of model parameters.

I would like to acknowledge the feedback and contributions of users and developers of the CRYSTALS software – in particular Pascal Parois, Bruce Foxman and David Watkin – which have led to investigations of many of the refinement features described in this chapter.

References

1. Brown PJ, Fox AG, Maslen EN, Keefe MA, Willis BTM (2006) Intensity of diffracted intensities. In: International tables for crystallography. International Union of Crystallography, pp 554–595
2. Sanjuan-Szklarz WF, Hoser AA, Gutmann M, Madsen AØ, Woźniak K (2016) Yes, one can obtain better quality structures from routine X-ray data collection. IUCrJ 3:61–70
3. Sheldrick GM (2015) Crystal structure refinement with SHELXL. Acta Crystallogr Sect C Struct Chem 71:3–8
4. Wilson AJC (1976) Statistical bias in least-squares refinement. Acta Crystallogr Sect A 32:994–996
5. Dominiak PM, Volkov A, Li X, Messerschmidt M, Coppens P (2007) A theoretical databank of transferable aspherical atoms and its application to electrostatic interaction energy calculations of macromolecules. J Chem Theory Comput 3:232–247
6. Fugel M et al (2017) Probing the accuracy and precision of hirshfeld atom refinement with HARt interfaced with Olex2. IUCrJ 5:32–44
7. Schröder L, Watkin DJ, Cousson A, Cooper RI, Paulus W (2004) CRYSTALS enhancements: refinement of atoms continuously disordered along a line, on a ring or on the surface of a sphere. J Appl Crystallogr. https://doi.org/10.1107/s0021889804009847
8. King MV, Lipscomb WN (1950) The X-ray scattering from a hindered rotator. Acta Crystallogr 3:155–158
9. Reilly AM, Morrison CA, Rankin DWH, McLean KR (2011) Using molecular-dynamics simulations to understand and improve the treatment of anharmonic vibrations. II. Developing and assessing new Debye-Waller factors. Acta Crystallogr Sect A Found Crystallogr 67:346–356
10. Betteridge PW, Carruthers JR, Cooper RI, Prout K, Watkin DJ (2003) CRYSTALS version 12: software for guided crystal structure analysis. J Appl Crystallogr. https://doi.org/10.1107/s0021889803021800
11. Binns J et al (2017) Phase transition sequences in tetramethylammonium tetrachlorometallates by X-ray diffraction and spectroscopic measurements. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 73:844–855
12. Bailey PJ et al (2010) A new synthesis of charge-neutral tris-pyrazolyl and -methimazolyl borate ligands. Chem Eur J 16:2819–2829
13. Barrett ES, Irwin JL, Edwards AJ, Sherburn MS (2004) Superbowl container molecules. J Am Chem Soc 126:1
14. Anderson E, Bai Z, Bischof J, Blackford S, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Sorensen D (1999) LAPACK users' guide. Society for Industrial and Applied Mathematics, Philadelphia
15. Wang Q, Zhang X, Zhang Y, Yi Q (2013) AUGEM: automatically generate high performance dense linear algebra kernels on x86 CPUs. In: International conference for high performance

computing, networking, storage and analysis, SC, IEEE Computer Society. https://doi.org/10.1145/2503210.2503219
16. Wang E et al (2014) Intel math kernel library. In: High-performance computing on the Intel® Xeon PhiTM. Springer, Berlin, pp 167–188
17. Parois P, Cooper RI, Thompson AL (2015) Crystal structures of increasingly large molecules: meeting the challenges with CRYSTALS software. Chem Cent J. https://doi.org/10.1186/s13065-015-0105-4
18. Spek AL (2015) Platon squeeze: a tool for the calculation of the disordered solvent contribution to the calculated structure factors. Acta Crystallogr Sect C Struct Chem 71:9–18
19. Frigo M, Johnson SG (2005) The design and implementation of FFTW3. In: Proceedings of the IEEE, pp 216–231
20. Coster D, Knol KS, Prins JA (1930) Unterschiede in der Intensität der Röntgenstrahlen-reflexion an den beiden 111-Flächen der Zinkblende. Z Phys 63:345–369
21. Hamilton WC (1965) Significance tests on the crystallographic R factor. Acta Crystallogr 18:502–510
22. Rogers D (1981) On the application of Hamilton's ratio test to the assignment of absolute configuration and an alternative test. Acta Crystallogr Sect A 37:734–741
23. Flack HD (1983) On enantiomorph-polarity estimation. Acta Crystallogr Sect A 39:876–881
24. Flack HD, Bernardinelli G (2000) Reporting and evaluating absolute-structure and absolute-configuration determinations. J Appl Crystallogr 33:1143–1148
25. Cooper RI, Watkin DJ, Flack HD (2016) Absolute structure determination using CRYSTALS. Acta Crystallogr Sect C Struct Chem 72:261–267
26. Parsons S, Flack HD, Wagner T (2013) Use of intensity quotients and differences in absolute structure refinement. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 69:249–259
27. Cooper RI, Flack HD, Watkin DJ (2017) HUG and SQUEEZE: using CRYSTALS to incorporate resonant scattering in the SQUEEZE structure-factor contributions to determine absolute structure. Acta Crystallogr Sect C Struct Chem 73:845–853
28. Brázda P, Palatinus L, Babor M (2019) Electron diffraction determines molecular absolute configuration in a pharmaceutical nanocrystal. Science (80-.) 364:667–669
29. Waser J (1963) Least-squares refinement with subsidiary conditions. Acta Crystallogr 16:1091–1094
30. Hirshfeld FL (1976) Can X-ray data distinguish bonding effects from vibrational smearing? Acta Crystallogr Sect A 32:239–244
31. Thorn A, Dittrich B, Sheldrick GM (2012) Enhanced rigid-bond restraints. Acta Crystallogr Sect A Found Crystallogr 68:448–451
32. Parois P, Arnold J, Cooper R (2018) An enhanced set of displacement parameter restraints in CRYSTALS. J Appl Crystallogr 51:1059–1068
33. Schomaker V, Trueblood KN (1968) On the rigid-body motion of molecules in crystals. Acta Crystallogr Sect B Struct Crystallogr Cryst Chem 24:63–76
34. Madsen AØ (2006) SHADE web server for estimation of hydrogen anisotropic displacement parameters. J Appl Crystallogr 39:757–758
35. Marsh RE, Bernal I (1995) More space-group changes. Acta Crystallogr Sect B Struct Sci 51:300–307
36. Marsh RE (2005) Space group P1: an update. Acta Crystallogr Sect B Struct Sci 61:359
37. Harlow RL (1996) Troublesome crystal structures. Prevention, detection, and resolution. J Res Natl Inst Stand Technol 101:327
38. Schwalbe CH (2018) Should we remediate small molecule structures? If so, who should do it? Crystallogr Rev 24:217–235
39. Spek AL (2020) CheckCIF validation ALERTS: what they mean and how to respond. Acta Crystallogr Sect E Crystallogr Commun 76:1–11
40. Harrison WTA, Simpson J, Weil M (2010) Acta Crystallographica section E: structure reports online: editorial. Acta Crystallogr Sect E Struct Rep 66:1–2

41. Henn J, Meindl K (2015) Statistical tests against systematic errors in data sets based on the equality of residual means and variances from control samples: theory and applications. Acta Crystallogr Sect A Found Adv 71:203–211
42. Henn J, Meindl K (2014) About systematic errors in charge-density studies. Acta Crystallogr Sect A Found Adv 70:248–256
43. Meindl K, Henn J (2008) Foundations of residual-density analysis. Acta Crystallogr Sect A Found Crystallogr 64:404–418
44. Prince E (2004) Mathematical techniques in crystallography and materials science. Springer, Berlin
45. Prince E, Nicholson WL (1985) In: Wilson AJC (ed) Structure & statistics in crystallography: proceedings of the Symposium on Crystallographic Statistics, held in Hamburg, West Germany in August, 1984 in the course of the Thirteenth International Congress of the International Union of Crystallography, A. J. C. Adenine Press
46. Wilson CC, Morrison CA (2002) Structural and theoretical investigations of short hydrogen bonds: neutron diffraction and plane-wave DFT calculations of urea-phosphoric acid. Chem Phys Lett 362:85–89
47. Reilly AM et al (2016) Report on the sixth blind test of organic crystal structure prediction methods. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater. https://doi.org/10.1107/S2052520616007447
48. Deringer VL et al (2014) Ab initio ORTEP drawings: a case study of N-based molecular crystals with different chemical nature. CrystEngComm 16:10907–10915
49. Van De Streek J, Neumann MA (2010) Validation of experimental molecular crystal structures with dispersion-corrected density functional theory calculations. Acta Crystallogr Sect B Struct Sci 66:544–558
50. Kyriacou A et al (2013) Combined X-ray and neutron diffraction Rietveld refinement in iron-substituted nano-hydroxyapatite. J Mater Sci 48:3535–3545
51. Liu H et al (2017) Sensitivity and limitations of structures from X-ray and neutron-based diffraction analyses of transition metal oxide lithium-battery electrodes. J Electrochem Soc 164:A1802–A1811
52. Il Kim Y, Jeon MK (2004) Combined structural refinement of Bi4Ti3O 12 using X-ray and neutron powder diffraction data. Mater Lett 58:1889–1893
53. Petrícek V, Dušek M, Palatinus L (2014) Crystallographic computing system JANA2006: general features. Zeitschrift fur Krist 229:345–352
54. Coelho AA (2018) TOPAS and TOPAS-academic: an optimization program integrating computer algebra and crystallographic objects written in C++: an. J Appl Crystallogr 51:210–218
55. Toby BH, Von Dreele RB (2013) GSAS-II: the genesis of a modern open-source all purpose crystallography software package. J Appl Crystallogr 46:544–549
56. Grzechnik A, Meven M, Paulmann C, Friese K (2020) Combined X-ray and neutron single-crystal diffraction in diamond anvil cells. J Appl Crystallogr 53:9–14

Struct Bond (2020) 185: 69–140
https://doi.org/10.1007/430_2020_63

Published online: 8 October 2020

Leading Edge Chemical Crystallography Service Provision and Its Impact on Crystallographic Data Science in the Twenty-First Century

Simon J. Coles, David R. Allan, Christine M. Beavers, Simon J. Teat, Stephen J. W. Holgate, and Clare A. Tovee

Contents

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/430_2020_63) contains supplementary material, which is available to authorized users.

S. J. Coles (✉)
School of Chemistry, Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, UK
e-mail: s.j.coles@soton.ac.uk

D. R. Allan and C. M. Beavers
Physical Sciences, Diamond Light Source, Harwell Science and Innovation Campus Didcot, Oxfordshire, UK
e-mail: david.allan@diamond.ac.uk; christine.beavers@diamond.ac.uk

S. J. Teat
Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
e-mail: sjteat@lbl.gov

S. J. W. Holgate and C. A. Tovee
Cambridge Crystallographic Data Centre, Cambridge, UK
e-mail: holgate@ccdc.cam.ac.uk; tovee@ccdc.cam.ac.uk

Abstract National facilities provide state-of-the-art crystallographic instrumentation and processes and tend to act as an indicator for the direction of a community in the medium term. There has been a significant step up in terms of instrumentation and approach in the last 10 years which has driven data generation. This has had a significant impact on databases – in turn we observe a substantial change in the use of the Cambridge Structural Database (CSD) from relatively basic search/retrieve to gaining deep understanding about factors that govern the solid state. Databases are now able to drive new science in areas such as crystal engineering. Looking forward, we will see more automated pipelining of the data generation process, and this will require better integration with databases. Databases will provide more predictive power – and this will inform the science/crystallography that should be done.

Keywords Cambridge Structural Database (CSD) · Central facilities · Crystal structure data · Crystallographic instrumentation · Crystallography · Data science · Single-crystal X-ray diffraction · Structural informatics · Synchrotron

1 Introduction to the Modern Crystallographic Environment

1.1 The Development of the Crystallographic 'Facility'

As the crystallographic technique matured during the twentieth century, the instrumentation and nature of the laboratory evolved. In the first half of the century, the laboratory tended to consist of bespoke equipment, often constructed in-house and only used by a small number of highly trained scientists whose research area was centred around solving key crystal structures. However, the real power of the technique was demonstrated around the middle of the century as much larger protein structures began to be determined [1]. This scientific development, along with the

greater general demand for crystal structures, fuelled the technological development of instrumentation capable of providing a higher throughput, but still relied on manual operation. This phase of technical progression powered significant developments in the 1960s and 1970s which saw the birth of 'chemical crystallography' as a distinct subject area. At the time, the main driver in chemical crystallography was the demand to establish the identity of synthesised compounds, and perhaps the most significant contribution to this came from a burgeoning organometallic chemistry community that had a need to characterise structural features and bonding in the vast number of exciting new compounds they produced. At this time, a crystallography laboratory would produce about one structure a month (about 500 reflections could be measured in a working day, and hence the size of the structure also had a big influence on data collection time). During the 1970s, instrumentation moved away from manual operation and became increasingly computer-automated, and, also, access to computing facilities and software for structure refinement became more available. The laboratory throughput rose to around 2–3 structures per week. During the 1980s, in-house mainframe computing became accessible, generating another increase in throughput and the transition to a service culture – however, the output of the laboratory essentially depended on how many diffractometers were available! The 1990s saw the (re)introduction of area detection methods and low-cost personal computing, and by the middle of this decade, charge-coupled device (CCD) detectors prevailed. These instruments could readily produce two datasets in a 24-h period. Also at this time, synchrotron-based instrumentation became available for chemical crystallography [2]. These facilities were still largely based on traditional laboratory equipment, but they provided the ability to look at smaller and more weakly diffracting crystals – this considerably expanded the range of science addressable by the technique.

Chemical crystallography became a thriving discipline as a result of these innovations. For example, the notion of crystal engineering was born and quickly grasped by a range of communities going far beyond the realm of the individual crystallography group. Crystal engineers are not the only example of the adoption of chemical crystallography by modern chemistry fields – the rapid growth of supramolecular chemistry and metal-organic framework design has been based in part on the revealing nature of crystal structure results. Furthermore the service crystallography culture has also become thoroughly embedded as a popular tool for the synthetic chemist, in part due to the unambiguous nature of a crystal structure result, but also to the fact that in many cases the 'waiting time' to get a result has become comparable to that of commonly used spectroscopic methods.

Additionally, these advances in instrumentation have fuelled the rise of more advanced single-crystal techniques in chemistry, such as high pressure [3, 4], in situ process monitoring [5] and photocrystallography-based studies [6]. This review is concerned primarily with service crystallography, but it will touch on some of the implications that advances in this area have for these techniques.

This review begins by considering the notion of the modern 'chemical crystallography facility'. That is, driven by the nature and volume of samples examined, the practice of service chemical crystallography now generally needs to go beyond

simply installing a single instrument (and crystallographer!). The modern chemical crystallography service needs to be based on the integration of advanced instrumentation with people, software and processes. National facilities have existed for a long time and have invariably been the driver and proving-ground for both new technological advances and in terms of how a facility should be run. For example, in these facilities automation, both of hardware and software, has been developed. However, the throughput of these facilities meant that data management, that is, the handling of the high volumes produced by the instruments, processing it and organising the results, has also become an absolute necessity. We therefore consider 'the facility' to be a combination of all these technologies that serves a range of different users with different chemistry research and different crystallographic experience.

We will draw on the experiences of two different types of facility in order to illustrate the main advances raised in this article. These are world-leading, national-scale facilities, with the principle of this review being that they have a mandate to pioneer the technique and that the innovations developed by these facilities will in turn become common practice in all facilities. The obvious examples of this type of facility are at national laboratory synchrotron sources, and we will draw heavily on two main examples of these, (a) beamline I19 [7, 8] at the Diamond Light Source in the UK and (b) beamlines 11.3.1 and 12.2.1/2 [9] at the Advanced Light Source in the USA. Further to these, the National Crystallography Service (NCS) in the UK [10, 11] operates at the national scale as an Engineering and Physical Science Research Council-funded National Research Facility [12] but is university-based, and this provides an illustration of how the technologies developed at synchrotrons can be translated to the home laboratory.

The modern service crystallography facility is however only one half of the topic of this review. Chemical crystallography facilities worldwide now generate results at an unprecedented rate, and this has a knock-on effect – the data explosion [13]. Not only does the facility have to cope with this situation, but so does the dissemination process – publishing processes are struggling to deal with this, but still we have seen a dramatic growth of the crystallographic databases. The second half of this review looks at the rise of crystallographic data, which has powered development of crystallographic databases. As the most relevant database to the chemistry community, we mainly consider herein the growth of the Cambridge Structural Database, the 'CSD' [14]. The CSD recently accumulated its millionth structure [15], but the state of the art in crystallographic data science does not stop with merely aggregating crystal structures – data mining approaches enable the understanding of trends and development of general rules. Modern approaches and large volumes of data now enable the development of 'knowledgebases', and this review will illustrate how it is possible to go beyond crystallographic knowledge and use it to drive new science.

1.2 The State of the Art

This article is primarily concerned with the rapid pace of developments relating to the capability of service chemical crystallography and its outputs in the last 10 years. At first thought there are, therefore, two perspectives to consider – those of equipment and data. However, there is more involved than merely considering hardware/software technology and the volume of data produced.

Firstly, one must consider holistically the whole facility and the processes that occur within it. Herein we use the examples of a home-based national facility and the centralised synchrotron facility to illustrate the current state of the art and thereby indicate the future direction for conventional facilities. This is because the purpose of these laboratories is to undertake work that cannot be performed in a conventional facility, and they thereby experience the toughest challenges, which have to be addressed with the development of new technologies and procedures.

Secondly, it is of upmost importance to be aware that the data collections we have now accumulated are so authoritative and comprehensive that they in fact provide a driver for the structural science that is undertaken. Whole sub-disciplines are now predicated on the knowledge we have gleaned, e.g. on geometry and intermolecular interactions, from data mining these collections. Therefore, it is not uncommon for, e.g. metal-organic framework (MOF) design or crystal engineering groups to operate their own facilities or to dominate and drive local service chemical crystallography facilities.

1.2.1 The Home Laboratory Instrumentation

Recently, the nature and complexity of the products of synthesis in chemistry has increased significantly, while service crystallographers supporting this work are generally doing so with a mature technology developed in the 1990s. Furthermore, there have been numerous developments in the scientific applications of the technique, for example, dynamic crystallography [16], that pose significant challenges for this technology. This resulted in a demand for more advanced instrumentation and methods, which has been well met and accordingly drove the capabilities of the technique forward.

X-Ray Sources

It is now possible to generate relatively high flux X-rays in the home laboratory compared to the traditional sealed tube approach. Traditionally a graphite monochromator would have been used to select the appropriate wavelength radiation from a water-cooled X-ray tube system. However, X-ray focusing optics were developed for chemical crystallography [17] in the early part of this century. This step change was driven by the specific need for high-intensity molybdenum radiation and was

adapted from a technology used with copper radiation by the protein crystallography community. These were pioneered for rotating anodes, but very rapidly became integrated as standard for the follow-on technology from sealed tube systems [18]. The inclusion of focusing optics on all X-ray generator types for chemical crystallography produced an increase in X-ray flux of around 6–10 times across the board. Around this time microfocus sources became ubiquitous – these are essentially the same technology as the sealed tube but tuned so that the electron beam is highly focused into a small spot on the anode. This results in an X-ray beam that is considerably more focused and therefore has a greater flux density for the same amount of power.

For the rotating anode, there is now even more potential. Within the last decade, it has become possible to produce a range of focusing mirrors with different capabilities. In particular for their highest flux rotating anode, Rigaku Oxford Diffraction can produce different optics – the NCS has two diffractometers on the same X-ray source, one with a very high flux (VHF) mirror and the other with an ultra-high flux (UHF) version. The UHF optic provides a sharper focus of all the X-rays emitted from the rotating target, resulting in a considerable increase in flux density at the sample. The NCS UHF system has a beam with a 70 μm focus at the sample, while the VHF is 150 μm, with the former giving an intensity in the home laboratory that is of a similar order to that of a second-generation synchrotron single-crystal diffraction facility.

The key to producing high flux X-rays by electron bombardment of a metal target is the efficient dissipation of the heat generated by the process. This is achieved in the rotating anode by having a moving target, and a faster, larger, target will produce greater strength X-rays. However, new approaches through the use of liquid gallium and indium alloys [19] as a target are particularly promising for improving the dissipation of heat and thereby producing higher-powered X-rays. These sources use a liquid metal, which is injected into the path of the electron beam and thereby generates X-rays, while the metal is gathered, cooled and recycled for subsequent X-ray generation [20]. This type of system is now marketed by the company Excillum and has been applied to numerous X-ray diffraction, imaging and spectroscopy techniques. With wavelength of 1.35 Å and 0.51 Å for the gallium and indium sources, respectively, it is possible to approximately mimic the traditional silver (0.48 Å) or molybdenum (0.71 Å) source for chemical crystallography and that of the system of choice for larger unit cell systems, copper (1.54 Å). Comparison measurements show this type of source to be stronger than molybdenum-based rotating anode systems.

X-Ray Detectors

Early crystallographic laboratories depended on photographic film recordings of the diffraction data. This painstaking process was later augmented by systems which used photographs for indexing and point detectors for the actual intensity data collections – this combination provided the reciprocal space detail of a 2D detector

but the speed and improved counting statistics of an ion chamber or pin diode. The development of electronic area detectors follows this struggle – striving for the highest possible contribution to spatial detail and angular coverage while also producing statistically reproducible intensity measurements. A range of different technologies was investigated. Multiwire proportional chamber (MWPC) gas detectors were capable of detecting many simultaneous diffraction spots but had limited global count rates. They were superseded by the much more widely used image plate detectors and CCD detectors. The image plate detector, which is still in use, has an extremely large format with excellent resolution, low noise and high sensitivity. It is ideal for studies requiring long exposure times, but the duty cycle, particularly the readout time, is considerable for experiments requiring several images. In comparison CCD detectors have a similar pixel size to image plates but have a significantly faster duty cycle, and, although they have greater noise than image plates (both dark current and readout noise), the very short exposure times required for single-crystal diffraction mean that the background counts generated by the electronics on each exposure are negligible compared to the signal. With the advent of large-format CCD detectors, the use of image plate-based systems declined rapidly.

CCD detectors have been in common use for over 20 years and are considered by many to be reaching the technical limit of their capabilities. Arguably they provided a step change for the diffraction experiment upon their introduction, and the CCD detector, as known to X-ray crystallography, consisted most simply of a visible light-sensitive CCD sensor and an X-ray phosphor. Through the photoelectric effect, the visible light produces a charge in the pixels, and voltage changes push the charges on the pixels through neighbouring pixels out to a limited number of amplifier nodes. Here the charges are converted to an analogue voltage, which then needs translating into digital counts for the pixels. This readout process, where charges are shifted through rows of pixels, requires the sensor to be dark, i.e. no X-ray photons impinging on the sensor, to maintain the integrity of the intensity counts and positions [21]. Advances in chip manufacture lead to the development of CMOS chip-based detectors. The primary difference between CCD and CMOS [22] chips is that each CMOS pixel has its own analogue to digital converter, so that an extended dark period is not required for readout. The pixel readout is parallelised in CMOS detectors, with pixels read out sequentially by multiplexers dedicated to a row or region of the detector. With multiple multiplexers per detector, the readout of the entire detector can be extremely quick. Monolithic CMOS detectors available today, such as the Bruker PHOTON series, detect visible light created by an X-ray phosphor and offer high-speed data collections due to their ability to collect continuous shutterless rotation data with no count rate-based saturation.

Along with the evolution of detector chip architecture, sensor materials have also been evolving. Modern microchip fabrication techniques have allowed for another detector advance – the technique of bump-bonding which joins a sensor module to the CMOS chip with a microscopic solder bump. This technique allows the sensor material and the chip material to be decoupled and optimised separately. These Hybrid Photon Counting (HPC) detectors, originally developed for particle tracking (e.g. at CERN), allow for faster and more sensitive data collection. These detectors

have rapidly found application in many areas and have been very successfully adapted to detect X-rays and used extensively for imaging and diffraction experiments. A major advantage of this approach is that direct detection of photons is possible, which is in contrast to the CCD detector that has to convert impinging photons into light which is read out as an analogue signal and then transformed into a digital signal. In this section we give brief illustrations of how this new technology has been applied to university/home laboratory-based facilities. However, there is further detail in the following section on synchrotron instrumentation.

Over a decade ago, the company Dectris launched the revolutionary Pilatus detector [23], which took silicon HPC technology and applied it to the crystallographic technique, and it rapidly became the detection system of choice at macromolecular crystallography (MX) beamlines around the world. A modest number of these detectors were installed in laboratory-based facilities; however, in the last decade, this technology has moved on considerably and is now readily available for chemical crystallography and the home laboratory. This is illustrated by the fact that major diffractometer manufacturers now sell their own versions of this technology as an off-the-shelf package specifically for chemical crystallography. In fact, Rigaku Oxford Diffraction has even gone as far as only selling HPC-based detectors for their X-ray systems, and the HyPix-6000HE [24] is specifically designed for single-crystal diffraction. The HyPix detector is readily paired with a rotating anode generator, which, due to the low noise level and coupling of detection to readout electronics, is particularly good at measuring diffraction from small and poorly diffracting crystals. Alternatively, Bruker has developed the PHOTON series of detectors [25], based on CMOS technology [26] which is a single monolithic charge integrating pixel array which can simultaneously count and integrate incoming photons and therefore has no count rate saturation point.

The combination of high flux sources and fast, noiseless detectors means that copper-based systems are now much more viable for chemical crystallography in the home laboratory. Due to the longer wavelength of copper radiation, it is necessary to position the detector in many different locations to cover a suitable volume of reciprocal space, and therefore a faster detector alleviates this time-consuming problem. This combination is very powerful, due to the much higher brilliance of copper radiation over the traditional molybdenum target.

1.2.2 Synchrotron Instrumentation

Sources

A ‘synchrotron’ is a particle accelerator, which is in fact something of a misnomer for this type of light source as the X-ray producing component is the storage ring and as such the particles are not accelerated and rather their energy is just maintained. Light sources generally use electrons, but there are a few that use positrons, e.g. PETRA-III in Hamburg. The electrons are generated in an electron gun and then accelerated in a linear accelerator before being passed into the booster ring,

which is the actual synchrotron component, that accelerates them to the full energy of the storage ring. At this point the electrons are travelling at around 99.9999% the speed of light, with an effective mass similar to that of a proton, and they can then be passed into the storage ring. Unlike a laboratory source, the X-rays are not generated by colliding with a target, but by making the electrons deviate from their accelerated trajectory, and as they do so, they have to give off a photon to conserve angular momentum. The energy of the photon depends on the energy of the electron and the radius of the turn. The lost energy is replenished using a radio-frequency (RF) cavity. Due to the combination of bends and straights, the storage ring's magnetic lattice forms more of a polygonal shape than a circle.

Synchrotron light sources have undergone an evolution, from particle physics playthings to large-scale user facilities over the past 70 years. Initially only used for accelerating particles for collision experiments, the observation of visible light emanating from a glass vacuum vessel on the 70 meV synchrotron at General Electric's Schenectady, NY, labs [27] changed the trajectory of large-scale user facilities entirely. Within years, scientists came to these particle accelerating sites and harvested the photons produced by particle physics storage ring operations – this is considered the first generation of synchrotron light sources, aka, parasitic operation. Beamlines, which direct and condition the light, were built in a generic fashion – the instrumentation attached was changeable and ad hoc. The second generation of light sources are those which were designed for the production of light for scientists, and they were built specifically to produce light in certain regions of the electromagnetic spectrum [28]. The macromolecular crystallography community were early adopters of synchrotron radiation as the highly collimated beams and high photon fluxes offered by second-generation machines provided significant advantages over laboratory-based instrumentation. The beamlines on these early synchrotrons were often taking light from bending magnets, which produced a smoothly varying energy spectrum. Bending magnets are characterised by their critical energy, which is where the total emitted light from the bending magnet can be divided in half. In the equations below [29, 30], it is shown that this can be simplified to a relationship between the storage ring energy and the bending magnet field strength.

$$E_c = \frac{3e\hbar B\gamma^2}{2m} \quad \gamma = \frac{E_e}{mc^2} \tag{1}$$

$$E_c = 0.665 E_e B$$

Equation 1 The relationship between electron beam energy (E_e, in GeV), magnetic field strength of a bending magnet (B in T) and the critical energy of the resulting radiation (E_c in keV). Due to e, ħ and m being constants, a simplified form has been produced.

These parameters, as well as the overall circumference of the storage ring, are carefully chosen to place the bending magnet critical energy in a suitable range for the experiments to be hosted at the synchrotron; this value was historically chosen

around 10 keV (1.24 Å) for diffraction studies. The light produced is in the form of a flattened cone at a tangent to the curved path of the electrons and with a fan angle equal to the angular change in the path of the electrons around the curve. As the critical energy is proportional to the magnetic field strength, higher photon energies can be achieved through the use of a 'superbend' where a conventional bending magnet (which will typically have a field of around 1 Tesla) is replaced by a superconducting magnet (which can have a strength in excess of 4–5 Tesla). In this manner, the radiated power output can be amplified by a factor of 20 or more with the critical energy increased by a multiple of about 5 with flux and brightness gains of around an order of magnitude [31, 32].

The transition from second-generation to third-generation synchrotron light sources was accomplished once storage rings were designed with long straight sections to accommodate insertion devices, which are alternating magnetic arrays placed above and below the electron beam. Some second-generation sources were later adapted to take 'insertion devices'; however, it was not until the advent of third-generation sources that most X-ray diffraction beamlines took advantage of insertion devices, which come in two main flavours: wigglers and undulators. All insertion devices manipulate the electron beam so that it oscillates in the horizontal plane by being placed through a periodically alternating magnetic field which is aligned orthogonally to the path of the beam. If the oscillating excursions of the electron beam are larger than the radiation cones emitted from each 'wiggle', then the intensities of each period in the oscillation are added, and the device is called a wiggler. The magnets in a wiggler are often superconducting with field strengths on the order of 2–5 Tesla. The resulting spectrum is smoothly varying with a high characteristic energy and with a photon flux commensurate with the number of magnetic poles used in the device.

In an undulator, the magnetic fields are smaller; hence, the excursions are much gentler; the radiation cones emitted from each oscillation overlap in such a manner that there is an interference effect. In this instance, the amplitudes are added (taking into account the phase difference from each contributing cone), and the sum for each contribution, or period, along the undulator is squared to produce the intensity. The intensity peaks at energies where the interference is constructive to produce a spectrum that is characterised by relatively sharp harmonic peaks. The magnets, usually powerful permanent magnets (with a strength of about 1 Tesla), are held in jaws placed parallel either side of the electron beam. The energies of the harmonics can be tuned by varying the gap between the jaws (that is the strength of the magnetic field felt by the electron beam) so that a range of energies are available by a combination of adjusting the undulator gap and hopping to a different harmonic. As the oscillation of the electron beam within the undulator is relatively shallow, the X-ray beam has an extremely small divergence in comparison to either a bending magnet or a wiggler.

With the advent of third-generation synchrotron sources, crystallography beamlines now tend to be based on undulator sources, as the low divergence, high flux X-ray beams are ideal for the monochromatic methods of diffraction. The continuous, smooth, spectrum of a bending magnet or wiggler source can also be

well adapted for crystallography, especially when element-specific resonant data or simultaneous spectroscopic data is needed or if higher energies are required to penetrate in situ apparatus. In moving from second- through to third-generation synchrotrons, the source size has continually been reducing, due to accelerator physics improvements as well as the choice of insertion device becoming more commonly an undulator rather than higher divergence wigglers and bending magnets. Currently, most synchrotrons worldwide are considering or implementing low electron emittance upgrades, also known as brightness limited upgrades. This movement has centred around reducing the electron beam emittance; the primary way of achieving this is to reduce the angular bend per magnet by replacing one bending magnet with many lower field magnets. Variants on the multi-bend achromat storage ring design have been implemented in MAX-IV [33], will be implemented on SIRIUS and will be integrated into many existing facilities as upgrades, e.g. ESRF-EBS [34], APS-U, ALS-U and PETRA-IV. The choice of sources possible for crystallography beamlines will change in this fourth-generation, high-brightness, low-emittance regime: bending magnet sources will rarely be feasible for crystallography, wigglers will be chosen if a smooth continuum is necessary, and undulators will yield even higher brightness and flux. With source size reduction, it becomes easier to focus the beam to a smaller spot at the sample and thereby increase the flux density. Sub-micron spots will be routinely possible; however, when operating a rotation-based data collection approach in this regime, the limitations will be the speed and accuracy of the diffractometer and the ability to centre the sample.

Optics

For studies requiring a single well-defined wavelength, the energy is usually selected by a monochromator, which exploits the Bragg diffraction from crystals to select the required energy and direct the monochromatic beam to the downstream optics. The design of synchrotron monochromators varies considerably, but usually they involve a pair of matched crystals, a so-called double-crystal monochromator (DCM). These are often silicon polished to the (1 1 1) Bragg plane and, with the first crystal held at an angle to the primary beam, to select the energy, while the second crystal is positioned parallel to the first so that the monochromated beam is directed parallel to its original path, with a spatial separation from the main, unmonochromated beam. It is important to match the vertical beam divergence to the natural rocking width of the silicon reflection to maximise the flux – in higher divergence sources, a mirror may be used upstream of the monochromator to collimate the incoming X-rays, but the construction of this mirror imposes a restriction on the maximum photon energy of the beamline. The typical bandpass of a monochromator with this configuration is 0.1%, which is a very thin slice of the energy available from the source. As the angle of the first crystal is changed to select a specific energy, the second single crystal can be moved both horizontally and vertically to maintain the position of the X-ray beam leaving the monochromator – providing a 'fixed exit monochromator'. The huge

advantage of this configuration is that the following focusing optics and diffractometer do not have to be moved when the energy is changed.

The monochromators at crystallography beamlines in the future may follow some of the same principles of the classic DCM, or they might not even exist. The flux limiting optical element on most crystallography beamlines is the monochromator – three orders of magnitude can be lost in a silicon DCM. This is due to the extremely narrow rocking curve, which has been advantageous for spectral purity; however, this narrow level of bandpass is not necessary for most crystallographic experiments. The humble sealed tube has a much lower energy resolution, usually quantified as E/ΔE, equalling 150, whereas a typical Si DCM is closer to 7,000. Reducing this resolution to a more moderate level will not harm the crystallographic experiment, but it will allow access to unprecedented flux levels. There are various methods for increasing the bandpass of a traditional DCM – replacing the silicon crystals with multilayer mirrors can give E/ΔE values similar to those of a sealed tube. Bending the silicon crystals will also increase the bandpass. The most radical method for decreasing the flux loss at the monochromator is to eliminate it completely. The undulator harmonics on fourth-generation rings are expected to become more Lorentzian/Gaussian in shape, and it is likely that the harmonic peak will be suitably narrow not to need further monochromation. The harmonics can be dispersed spatially using a prism made from aluminium, or some other suitable material, and the required harmonic selected with a set of slits (the harmonic separator concept [35]).

Downstream of the monochromator there are usually mirrors to focus the beam in both the horizontal and the vertical planes to best match the sample size. Again, the design and configuration of the focusing optics can differ considerably between beamlines with the type and number of focusing elements varying to provide the desired beam size. This is commonly performed with either Kirkpatrick-Baez mirrors [36], a so-called KB pair, or a toroidal mirror. The KB pair consists of two mirrors that are initially flat, but then bent along their length and that are arranged horizontally and vertically to focus the beam in those planes. The toroidal mirror is a cylindrical mirror that is bent along the beam path to focus both horizontally and vertically. All mirrors, however, work at a grazing incidence angle (usually with an incidence angle of only a few milliradians), and they have coatings to supress the reflectivity of the higher energy harmonics from the monochromator (for Si (1 1 1) the strongest harmonic arises from the λ/3 component, associated with the (3 3 3) reflection, but it is also possible to observe the 'forbidden' (2 2 2) reflections, which manifests itself as a very weak λ/2 harmonic, for very strongly scattering samples). This would be detrimental in that resulting data will be contaminated by these harmonics and could produce effects such as unit cell doubling and tripling. The mirror coating needs to be selected to suppress both of these harmonics. As the mirrors are required to suppress harmonics over a large energy range (typically 5 keV to 35 keV), they are often prepared with parallel lanes of different coating materials, with each lane being used in a specific energy window. The coating materials have different atomic weights, e.g. Si, Cr or Pt, and by choosing

appropriately, the energy of a particular harmonic will not be passed through, e.g. at 8 keV, a Si coating with not allow 16 and 24 keV to pass.

As X-ray mirrors need to operate at grazing incidence, great care must be taken to ensure that their surfaces are free from errors in their figuring (slope error) that can result in texture in the profile of the focused beam. Additionally, the temporal beam stability can be affected by any vibration transmitted through the mirror mount – which can be greatly amplified due to the long optical lever from the source through the optical components to the sample. These issues are especially important for experiments requiring focusing on the order of 10 μm or less. A solution for these issues is the use of X-ray compound refractive lenses, which focus the beam through a process of refraction rather than specular reflection. As the refractive index for X-rays is close to and slightly less than 1, the lenses need to be concave to achieve focus rather than the convex form used for lenses to focus visible light. Differing focal lengths can be achieved by stacking a series of identical lenses into long rows, called a transfocator, with a different number of lenses used to bring a specified energy into focus at the same point. Compound refractive lenses produce extremely well-focused beams, with very little profile texture at ideal focus, and they are very tolerant of vibration, as they do not couple to the optical lever to the same degree as mirrors. There is some absorption by the lens material itself, as the X-ray beam must pass through it, so the choice of material is critical.

The advantages of increasing X-ray flux, using the above methods, are clear. However, there is a potential downside which the macromolecular crystallography community have had to tackle since the beginnings of structure solution from single-crystal diffraction at synchrotrons – that of the propensity for crystal decay in the X-ray beam. In the early days of synchrotron chemical crystallography, this was not an observable problem, due to chemical crystals being less prone to decay, the relative strength of a second-generation synchrotron and the sensitivity of the detection. However, with the technological developments since then, this effect is now far more likely and regularly being observed. This phenomenon is discussed further in Sect. 1.3.

End Stations: Detectors

The development of single-crystal diffraction techniques at synchrotron sources for atomic-resolution structure determination has been pioneered by the protein crystallography community, where the advantages of synchrotron radiation over lab sources were evident on early first- and second-generation sources [37]. As well as driving the development of sources and optics, protein crystallography was also instrumental in the advancement, or adaptation, of detector technology for the recording of the diffraction data. It is critical to remember, however, that the needs of a protein crystallography beamline and one optimised for smaller molecules are not identical. The push for large-format detectors largely stems from the d-space resolution necessary for protein crystallography at an X-ray energy which resolves the closely packed reflections from large unit cells. Protein diffraction data has a

smaller spread of intensities, due to the majority of the scattering atoms being close in atomic number. The mosaicity of a protein crystal is generally higher, and the diffraction limit is generally lower than that of a similarly sized crystal of a chemical compound. These differences are key in understanding how detectors were developed for protein crystallography and are also informative in how chemical crystallography at synchrotrons developed in parallel. It could be considered that the development of detectors suitable for this work presented the greatest initial challenge and it is only comparatively recently that detector technology is close to catching up with the full potential that third-generation synchrotron sources have to offer for crystallographic studies using single-crystal methods.

Dectris have developed a new series of pixel array photon counting detectors, Eiger, which have significantly smaller pixel sizes (75 × 75 μm rather than 172 × 172 μm) and with a much faster 3 μs operating at a frame rate of up to 3 kHz. With this new breed of detector, it is now possible to perform serial crystallography with wedges of data collected on several sample crystals per second from a slurry of microcrystals spread on a sample plate. At a synchrotron with small molecule crystals, there is the issue of encountering the limitations of the Dectris HPC chip designs, in the Eiger and Pilatus detectors. Because of the thresholding which is used to detect a single photon, a photon on the edge of two or more pixels may not create enough charge in any one pixel to be detected. Because of the relative sharpness of small molecule crystal peaks, this effect is more likely to affect small molecule data. This issue is avoided in HPC detectors built on the MediPix chip design [38], such as the X-Spectrum Lambda series, which allows adjacent pixels to essentially compare and account for these below threshold events. Another issue with all HPCs which is more commonly encountered in chemical and mineral crystals is the pileup effect. Essentially only one photon can be detected at a time in an HPC, and this creates a maximum photon rate that can be handled. Recently a workaround has been implemented in the Pilatus 3 and Eiger detectors, but this has only increased the limiting rate; it has not eliminated it. This pileup effect can lead to intense low angle peaks being underestimated. Alternatively, Bruker has developed the PHOTON series of detectors, based on CMOS technology [26] coupled to a phosphor. This detector is a single monolithic charge integrating pixel array which can simultaneously count and integrate incoming photons and therefore has no count rate saturation point – it is limited by the pixel well depth and readout time.

Looking forward, synchrotron beamlines of all types will certainly be working with brighter and more intense sources and will have to learn from the detectors of the FELs [39, 40]. Photon counting is generally too slow for FELs – most FEL detectors are integrating detectors. The construction is similar to HPCs, but the readout chip architecture is designed to measure the charge on the sensor pixel, not the arrival of individual photons. One of the current integrating detectors, like the JUNGFRAU [41] from PSI at the SLS, has been shown to collect high-quality protein crystallography data. The pileup issues seen in today's detectors will not be tolerable on future crystallography beamlines, when multiple orders of magnitude more flux is arriving at the sample.

End Stations: Small Molecule Specifics

The huge potential for using synchrotron radiation for the determination of small molecule structures, from crystals of a size that would formerly have been considered as single powder grains, was demonstrated by Harding in 1994 [42] with the structure solution of the industrially important catalyst precursor material aurichalcite. They used the SRS wiggler protein crystallography station, 9.6, at the Daresbury Laboratory with an Enraf-Nonius FAST (electronic area detector) diffractometer (0.4° frames, exposure time 40 s per frame, covering 200° in φ with the synchrotron ring current 20 mA (single-bunch mode)). Despite the low single-bunch ring current and the very small crystal size (100 × 40 × 5 μm), they obtained excellent results which demonstrated the need, and likely success, of a dedicated small molecule single-crystal diffraction station at SRS. The work, in part, was a precursor to the first beamline built at a synchrotron that was designed solely for the support of small molecule crystallography, station 9.8 at SRS Daresbury [43], which was built around a turn-key laboratory-based diffractometer system that was adapted for the synchrotron environment.

Thereafter a number of dedicated small molecule crystallography beamlines at synchrotrons also started out using commercial diffractometer and area detector systems. This approach had the strong advantage that while almost all users had no synchrotron experience, they were familiar with the software and operation of the diffractometer system. Furthermore, the hardware and software were well tried and tested by both the manufacturers and their customer base, with issues being found, reported and fixed quickly. This combination of factors meant that a largely unfamiliar chemical crystallography community was able to rapidly and confidently embrace the use of synchrotron radiation and quickly realise its benefits.

However, the rise of the third-generation synchrotrons with smaller source sizes resulted in smaller, more intense beams, especially from undulator sources, and flaws in this approach started to become evident. The diffractometers and detectors were too slow to make the best use of the high intensity, resulting in the beam being attenuated to get the diffracted intensities back into a range where they can be useful. Furthermore, the smaller beams decrease the tolerance on the sphere of confusion of the instrument and the alignment of the crystal. The manual goniometer head is then no longer able to position the sample well enough, and mechanical wear becomes a problem, as they were designed for a few crystals a day and not the 30 or 40 a day at the synchrotron.

Not for the first time, the chemical crystallography community is looking to protein crystallography beamlines to solve many of these problems. Air-bearing diffractometers such as the SmarGON D6 [44] are capable of omega rotation at speeds of 180°/s, compared with the 1°/s performance of traditional commercial goniometers. Generally, the crystal positioning in all axes on this and the other bespoke goniometers now in operation is motorised. Furthermore, detectors such as the photon counting Dectris Eiger 1M are able to collect 3,000 frames per s, c.f. a Bruker PHOTONII at 2 frames per s. Combining these technologies means a standard dataset could take as little as 30 s.

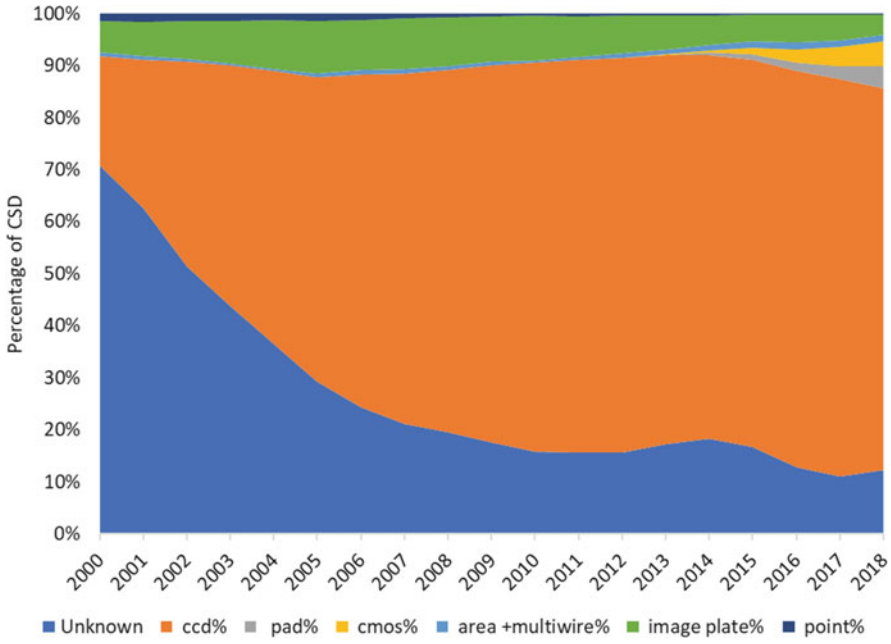


Fig. 1 Percentage of structures in the CSD collected using different detectors since the year 2000

Considering that interlocking the experimental hutch would take a minimum of 30 s and that crystal selection and mounting will take longer still, a good proportion of allocated beamtime would not be spent actually collecting data! This is clearly a waste of a precious resource, and again, the MX community has the answer – sample mounting robots. Crystals are pre-mounted, frozen and placed in a dewar for the robot to take and place on the diffractometer. The crystal is centred remotely by mouse-clicking on the centre of the crystal, which is then driven to a predetermined point in space. This now allows the whole experimental process to be run without the user being at the beamline – they mount their crystals in the comfort of their own labs, freeze and ship them to the synchrotron. With the crystals loaded into the dewar at the synchrotron by beamline staff, the experiment can be run by the users remotely. During the experiment the users will centre the crystal; assess the quality and intensity for the diffraction pattern; change attenuation or counting time as necessary; see how well it indexes; assess the data resolution; and determine the quality of the structure solution and refinement. I19 at Diamond Light Source (UK) is the first small molecule beamline to offer this capability to their users.

It is interesting to note the effect of advances in experimental practice and particularly instrumentation improvements on the data in the CSD. As mentioned above, modern X-ray detectors deliver greater accuracy and speed than those used in the early years of crystallography. The percentages of the different types of detector used to determine structures submitted to the CSD this century are shown in Fig. 1. As per the discussion above, the dominant use has been CCD detectors; however, in

the last 5 years, there has been a noticeable uptake in the use of Hybrid Photon Counting (CMOS and HPAD) detectors. The latter trend clearly points to this type of detector being the technology of choice for the coming years, and the discussion above indicates the effect that this will have on data quality.

The average collection temperature over the all the structures reported in the CSD has gone down approximately 15 K over the last 10 years from 204 K in 2009 to 189 K in 2018, although the range of temperatures has increased slightly as more experiments are performed to investigate the behaviour of materials under extreme conditions. The increased use of cooling technologies, particularly those using liquid nitrogen [45] to reduce thermal motion and possibly freeze out disorder, has undoubtedly played a role in this decrease.

1.2.3 The High-Throughput Process

The previous sections have considered hardware developments and illustrated their ability to provide very rapid and accurate data collection. However, these advances cannot be harnessed without development of the appropriate software to drive the process, manage it and deal with the output. Furthermore, to enable true high-throughput or optimum efficiency, this must happen in real-time in relation to the data collection process. Computing hardware can now be configured to process at high speed, and very powerful software can be written relatively easily, so it is not surprising that there have been software developments that are commensurate with the hardware improvements. Both national facilities and instrument manufacturers have dedicated teams to develop software to support and develop the experiment process, and this now involves functions beyond those of simply interfacing with the instrument and processing raw data. In this review we will focus only on those, more recent, aspects of software development that support higher-throughput data collection and the efficient operation of a facility, which predominantly falls into the categories of screening and automation.

Screening

As noted elsewhere in this article, the range of chemistry research being undertaken very often provides significant challenges for crystallography. Accordingly, samples can be poor in quality/crystallinity, and often it is not possible to recrystallise or grow better crystals. National facilities are generally more powerful than those available locally, so it is very common for them to be used to tackle the tougher problems that cannot be handled at home. Therefore, these facilities have become highly specialised to deal with this type of sample. The challenge with these samples is often the selection of the best possible crystal that will afford the best possible dataset, and so an efficient quality screening process is critical.

Due to the diversity of chemical samples, there has been little development of automated crystal growth, selection and mounting mechanisms in the way that

macromolecular crystallographers work. However chemical crystallography has a lot to benefit from by adopting fast diffraction screening. Traditionally this process has been one of taking a single diffraction snapshot and assessing quality by eye or by performing a short pre-experiment (typically 2 or 3 small scans) and judging the accuracy of the indexed unit cell. However, with state-of-the-art hardware coupled with well-developed software, new approaches are possible. It is now possible to automatically process, solve and refine data as it is being collected. Accordingly, the NCS screening procedure is now typically just to begin data collection and assess the structure as it progresses. This is made viable due to the fact that an HPC detector coupled with an intense X-ray source can collect data very quickly. Furthermore, part of the CrysAlisPro software [46], Autochem, provides the 'What is this?' routine, which automatically integrates, solves and refines the structure from the very beginning of a data collection. Typically, for a well-diffracting sample, an initial structure is available in around 2 min, and from this, it is possible to much better assess sample quality than by using traditional methods. For a challenging case, several samples will be trialled in this way and then the best one selected for data collection. This approach has not impacted the throughput of the facility, while at the same time it has considerably improved the quality of structure that one would typically get from such challenging samples.

On beamline I19 at Diamond, the software tool *screen19* has been developed to indicate whether count rate saturation has occurred for the Pilatus 2 M detector from a short, fine-sliced, data collection scan. It provides an estimate of the level of beam attenuation required to bring the count rate of the strongest reflections to below the required threshold based on an estimate of the mosaic width of the reflections. The program also indexes the observed reflections, to provide unit cell information, and provides a Wilson plot so that an estimate can be made of the maximum resolution that the sample diffracts to. The *screen19* program is launched automatically for each scan in any subsequent full data collection so that any degradation in the sample quality can be assessed. In parallel, the autoprocessing pipeline will provide a solved structure for each scan, or group of scans, used in the data collection, including the scan used for initial screening.

Remote Access

The use of remote access in small molecule crystallography is still very much in its infancy, and many of the techniques that have been implemented so far have been borrowed heavily from macromolecular crystallography, where remote access is now the exclusive mode of operation. In the paper of Johnson et al. [47], an excellent account is given of the development of the remote operation of beamline I19, at Diamond Light Source, for chemical crystallography studies. In their account, they have identified the following key areas:

Sample Loading and Transportation The preparation and mounting of samples have generally evolved in most chemical crystallography laboratories away from

traditional, glass-fibre, mounts to the use of magnetic bases fitted with Kapton micro-mounts. Crystals are usually bonded to the micro-mounts with a suitable inert oil and handled in a manner that depends on the specific properties of the sample, whether that is due to a possible chemical reaction or loss of solvent on exposure to the air. By immediately flash-freezing the crystals under liquid nitrogen in a Unipuck container, the samples can then be couriered to the synchrotron, ice-free, with the Unipucks held within the inert conditions of a dry shipping dewar. Each Unipuck can contain 16 sample mounts, and, with each dry dewar able to hold 7 Unipucks, a single shipment of 112 crystals can be made. On receipt of the dry shipper at the beamline, all 7 Unipucks can be loaded into the dewar of the robotic sample changer. Johnson et al. have found that distributing a number of crystals of the same sample in different pucks increases the chance of selecting a sample suitable for structure determination. This approach reduces the risk of losing the opportunity to collect data on a particular sample, in the event of an issue with an individual puck, as the spread of crystals between alternative pucks provides some redundancy.

Data Collection Remote data collection is carried out by the users connecting to the beamline using the NoMachine remote desktop software, and this can be carried out from the users' home laboratory (or indeed, with suitably fast broadband, from their home). The arrangement provides a very similar experience to the user being present in the control room of the beamline, although it is recommended that a workstation with at least two large screens is used due to the need to display several large software windows for General Data Acquisition (GDA), Information System for Protein Crystallography Beamlines (ISPyB) and some Experimental Physics and Industrial Control Systems (EPICS) controls. A key difference to the operation of the beamline GDA is the use of a software 'Baton', designating which computer currently holds control of the diffractometer and robot. This must be passed explicitly from the beamline to the user, with the local beamline staff able to take back control, via the baton, at any point. To aid communication between beamline staff and the remote users, a text messenger is incorporated within the GDA.

Data Processing With the more efficient use of the beamline and the corresponding increased rate at which data are recorded, the degree of data analysis that can be carried out during a session of beamtime can only be relatively rudimentary as there is time pressure to collect data from other crystals. The automated data processing pipeline can greatly assist in providing reduced data, and solved crystal structures, for well-behaved samples, but for more challenging cases, data analysis will require additional effort. For these more difficult datasets, which may involve twinning (or where there are multiple-crystal components) and major structural disorder, the diffraction images can be either imported directly into or converted into the appropriate file formats alternative processing packages that fit with the users' previous experience and preference. These software packages, whether they are public domain or proprietary, contain more bespoke tools for small molecule crystallography that allow useful integrated data to be produced from the usually very poorly diffracting crystals that are brought to the synchrotron.

Consequences of Remote Access More generally, remote access not only eliminates the need for users to travel to the synchrotron with the glass-ware containing their samples, which carries safety considerations, but it also offers them the potential to organise their beamtime into short, and more regularly spaced, intervals rather than whole day blocks separated by weeks. This more regular access provides a more timely turnaround for urgent, high-priority, samples. Remote access also allows greater cooperation between groups in a BAG (Block Allocation Group) as the GDA baton can be passed freely between geographically separated institutions in the same BAG so that each end user can run their own set of samples at a convenient period during the allocated beamtime. This greater flexibility is a compelling means of introducing new users.

Automation

Chemical crystallography facilities are now available for use by a wide range of researchers of varying experience working across a large range of scientific disciplines and often geographically distributed across the campus/region/country/world. Therefore, the ability to track, process and deliver a high volume of datasets is paramount and is now a critical capability for most service crystallography laboratories and a key component of being an efficient facility. Larger-scale facilities are developing information management systems to address this; for example, see ISPyB [48] described below. However, to a certain extent, every chemical crystallography facility requires systems that can support application, access, sample submission, sample tracking and reporting, data accessibility and data management. For a number of facilities, some of these may still be largely 'manual' processes; however, as rates and volumes of data production increase exponentially, these are beginning to become untenable. While crystallographic data is well structured and understood and only of medium size in terms of modern digital storage, over time the volume and heterogeneous nature of the files, ranging from binary images to small text (e.g. CIF) files, becomes an issue for data management at the facility scale. For example, as a high-throughput facility, the NCS diffraction laboratory at Southampton generates approximately 2,500–3,000 data collections per annum – this rate provides an indication of the issues that all laboratories will need to address in the future.

The recent development of highly automated methodologies for the production and purification of proteins, along with techniques for optimising crystallisation, has led to an ever-greater demand for the collection of X-ray diffraction data at MX beamlines at synchrotron sources. This greater demand has driven the automation of the beamlines with emphasis placed on intuitive and user-friendly software and the use of robotics for sample handling and exchange, with samples mounted in a variety of formats – including loops and X-ray diffraction-compatible crystallisation trays. With ever-brighter sources, coupled with high-precision goniometry and fast pixel array photon counting detectors, there has been a step change in data collection

efficiency with a commensurate increase in the volume of novel, and increasingly more challenging, protein structures that have been determined.

With the rapid development of sample preparation and data collection infrastructure, there has been a strong drive to manage the overwhelmingly large volumes of data that can be generated at each stage of the process: from initial sample production through to the finally realised crystal structure. To this end, many synchrotrons have developed a laboratory information system (LIMS). The LIMS system at Diamond, more widely known as ISPyB, provides a portal for users to track each stage of their experiment; detailed sample information can be added for each sample ahead of beamtime, along with the location of each sample within the shipping container. ISPyB can be used to generate the appropriate labels for the shipping containers to allow them to be couriered, and tracked, to the synchrotron. At the commencement of beamtime, ISPyB is updated with the specific location of each sample within the beamline's robotic sample changer. The location and all of the accompanying sample information can then be moved to the beamline's data collection, GDA, software from which each sample can be called and mounted for data collection. Usually data collection is carried out by the user accessing the beamline via remote login via their home institution. All of the raw and processed diffraction data is then linked to each sample along with all of the accompanying metadata, including the beamline parameters and images (snapshots) of the crystal from the beamline's sample viewing optics.

There has been a large shift for macromolecular crystallography to make LIMS, such as ISPyB, the core of the data collection, and data curation, process and to make the database itself the principal means of user interaction for all steps. For small molecule crystallography, progress in this area has been somewhat slower as there are few dedicated beamlines that exclusively support this technique and only a subset of these have any degree of robotic sample exchange. Beamline I19 at DLS, however, has facilities that are reminiscent of those on MX beamlines, including large-format pixel array photon counting detectors and a robotic sample changer, and is fully incorporated into the DLS GDA infrastructure. As a consequence, it is more fully compliant with the requirements for integration with ISPyB. Following the upgrade to the high-throughput diffractometer and robotic sample changer in experiments hutch 1 (EH1) of beamline I19 [49], there has been the adaptation of ISPyB to enable the recording of metadata specific to small molecule crystallography.

For MX, after initial data processing, the unit cell can be compared with similar cells for the user's current visit and previous visits, and the cell can also be searched for within the Protein Data Bank (PDB). For small molecule crystallography, a cell-matching search for similar unit cells in the user's current and previous visits can be undertaken, but, as yet, any search in the Cambridge Structural Database (CSD) needs to be undertaken outside the environment of SynchWeb with the CSD application ConQuest. It can be envisaged, however, that a suitable interface could be developed to pass cell information from SynchWeb to ConQuest or to provide a relatively seamless link between both applications.

Data Management, Retention and Availability

With such large volumes of data being generated, particularly via high-throughput and automated approaches, facilities are now having to seriously consider their position with respect to commitments to users and the broader community with respect to policies regarding the data they collect. Not only does the location of data have to be managed over time, but also its accessibility – and this applies to all scales of operation. National-level facilities, being centralised and multi-user operations, generally have data statements which illustrate the stance they have adopted. Beyond these statements, such facilities are in varying states of readiness with respect to having formal policies in place to address them. As the requirement to make a greater commitment to data management rises, the resourcing required is becoming more supported, and accordingly policies are becoming more widespread and meaningful. For example, the Diamond Light Source Data Management policy [50] statement declares 'The purpose of this policy is to provide Users conducting Peer Reviewed Research with information and guidance on Experimental Data ownership, storage, access and management and to ensure that Experimental Data is managed and used in ways that maximises public benefit'. Backing this up, the policy states that all data and metadata generated by facility users will be kept live for a minimum of 30 days after which a long-term archive copy is created – for the first 3 years, access to this is exclusive to the experiment team, but after this period, the data may be made openly available (under a CC-BY-4.0 licence). The ESRF data policy [51] has the same 3-year embargo period and commits to definite long-term storage of 5 years, but striving for 10 years. However, the ESRF policy goes further in that it makes a commitment to underpinning the publication process via provision of a DOI that can be cited in articles publishing the results derived from this raw data – this implies a much longer-term retention for data relating to publications. Conversely the Advanced Photon Source data retention policy [52] is rather sparse and makes no commitment to long-term storage, stating 'Experiment data and metadata collected at APS beamlines may be stored at and retrieved from the facility for at least 3 months'.

While such policies are generally the concern of central, national facilities, these commitments indicate the longer-term view and requirements of the research community. It will therefore become increasingly more important for smaller-scale facilities to consider their commitments to data management. This consideration is made increasingly more important by the imposed requirements of the funders and publishers of research for more transparent and rigorous access and management of data. Amongst the smaller-scale chemical crystallography facilities, there are varying attitudes and states of readiness towards these aspects of data management, and it will therefore take longer to address. It is clear that additional resources, such as finance and expertise, will need to be found in order for this to be implemented fully at any scale facility. It is also worthy of note that databases such as the PDB and CSD now facilitate and expose links, via DOIs, from records in their databases to raw diffraction data, which aids the discoverability of raw data.

1.3 Facilities and the Relationship Between Home Laboratory and Synchrotron

Nowadays many chemistry disciplines are utterly reliant on structural information to understand the complexity of a system or how it functions. These disciplines can produce very challenging samples for single-crystal diffraction, and also often it is not possible to recrystallise them or refine the synthesis method to generate larger crystals. Often these samples exceed the capability of local, traditional crystallographic facilities and recourse to larger-scale centralised facilities, and the appropriate expertise to handle such challenges is invaluable. To address these requirements of modern chemistry, the UK has a 'three-tiered' model (1 = 'home laboratory'; 2 = 'National Service laboratory'; 3 = 'synchrotron facility'). This means that weak diffractors can be screened on the most powerful laboratory source available before being referred to the synchrotron and ensures that the right samples are matched to the correct facility. This means that as we approach the diffraction limit for crystals of particular types of chemical systems, it is still possible to get structural information.

The suitability of a facility to particular samples is not only important for efficiency, but also we are now observing severe radiation damage for a significant proportion of samples investigated at a third-generation synchrotron [53]. In fact, it has been observed that samples previously considered to be 'radiation hard' are in fact affected by the radiation doses that modern synchrotrons generate. In some more extreme cases, there is clear evidence that the chemical makeup of a material can change as a result of exposure to this level of dose, i.e. the structure is not the same material at the end of the experiment as it was in the beginning. It is therefore now becoming necessary to adopt the approaches that the protein crystallography community have taken to address this problem [54] and the interplay between the home laboratory and synchrotron source will become crucial for successful data collection in certain cases.

A comparison between the two facilities several years ago [11] showed that a routine data collection with images collected at 1 s each on the attenuated synchrotron source is approximately equivalent to a 30 s per image data collection on the NCS state-of-the-art instrument at that time. Final *R*-factors for data merging and structure refinements were very comparable, so when taking into account the difference in time factor and attenuation applied, the synchrotron was collecting comparable data 100 times faster than the home laboratory. In the intervening time, both facilities have upgraded instrumentation – with the home laboratory installing a new HPC detector and the synchrotron upgrading the monochromator and goniometer and changing to a HPC detector.

1.3.1 Comparison Data Collection

We have therefore embarked on a comparison study, reported herein, with the intention of benchmarking these facilities relative to each other. This also provides a mechanism to describe the experimental procedure and make available template CIFs for a general data collection carried out on either facility by the NCS. Full procedural descriptions, along with data collection and refinement parameters and statistics, are provided as Electronic Supplementary Material to this article.

A suitable sample was selected, i.e. very stable and with suitable diffracting power to give workable results on both systems. Using a very accurate diode and taking into account the size of the beam and attenuation factors, it was measured and calculated that the flux density at the sample is over three orders of magnitude greater at the synchrotron source. Very similar data collection strategies were designed for both facilities in order to ensure a rigorous comparison with similar coverage. In order to obtain comparable data, the home laboratory collection was run at the equivalent of 100 s per degree, while the synchrotron collection was performed at 2 s per degree and using 10% of the incident beam. While the home laboratory experiments are clearly considerably slower, it can be seen that there is more observed data collected and the resulting agreement factors are better. One might now begin to conclude that for single crystals that are at the limit of hypothetical diffracting power, a very powerful home source will provide equally good results as a synchrotron. However, there will be samples where the hypothetical diffraction limit is not achievable in-house and the superior flux of the synchrotron will produce a better model. There will of course also be situations where the considerably more rapid data collection capability of the synchrotron will be of benefit. However one also has to be cautious as the dose received by the sample will often be very similar (less flux over a long time against more flux in a short period) in both facilities and therefore the best/correct strategy and facility must be selected in the case of samples that are prone to serious radiation damage.

1.4 Conclusion: Learning from Large Facilities

Large, national facilities drive innovation and act as a beacon for instrument and process development. These facilities indicate the future for conventional, smaller-scale facilities and act as an indicator of the direction in which the field is going in for the medium term. The example of the instrumentation and capabilities in the home laboratory of the NCS illustrates how the synchrotron-driven approaches can be translated to the everyday facility while demonstrating exactly what is possible and that the difference between local and national facilities needn't be so large.

There is a clear impact arising from the development of faster, more accurate technology and the increasingly automated methods that have been derived to make the most efficient and effect use of these instrument advances. Not only do these

developments lead to the generation of a greater volume of data, but they also enable examination on increasingly challenging samples. This provides us with the ability to examine extremely small and weakly diffracting crystals and increasingly is being used to drive the boundaries of structural science forward, e.g. the ability to collect data extremely rapidly enables the characterisation of dynamic processes in the solid state. However, the main, key point of this review is to demonstrate how these technological advances provide support for the challenges modern chemistry provides and most importantly is a significant driving force behind small molecule crystallography really becoming a more data-driven discipline.

2 The Impact of Technological Advances

2.1 The Communication of Results

Since communication in academia began, the scientific article has been viewed as the primary form of currency. It is worth taking a brief look at the history of the article to illustrate how it has changed over time. The first journal was the *Philosophical Transactions of the Royal Society* [55], started in 1665, which did include some 'data' and had an element of peer review. However, over the following centuries, there was a tendency for shorter, letter-type publications, which merely reported some final results. It was not until well into the twentieth century that the journal article was produced in the way we understand it today. It is interesting to note that as late as 1938, the journal *Science* relied wholly on personal solicitations for its content, while it was only in 1967 that the journal *Nature* introduced peer review – prior to which articles were primarily accepted based on editorial judgement. The following quote from Einstein relating to a *Physical Review*'s submission in 1936 is particularly revealing!

> We had sent you our manuscript for publication and had not authorized you to show it to specialists before it is printed. I see no reason to address the in any case erroneous comments of your anonymous expert. On the basis of this incident I prefer to publish the paper elsewhere

Peer review became ubiquitous as the method of quality control in the second half of the twentieth century not only due to the increasing specialisation of research but also due to the sheer volume being conducted and the fact that mass printing and distribution had become trivial. The result of these developments is the knock-on effect that the journal article has become the primary measure of the impact of a piece of work and now dictates much of the way the academic system operates. Research funding, journal quality, institutional esteem, personal esteem and promotion prospects are all affected by the perceived quality and volume of publication. Furthermore, the impact of work is judged on not only the quality of the journal but also a whole range of other metrics, such as those based on the number of citations an article receives. During this period the inclusion of crystal structures in articles has

massively increased and is well known to have an effect on how publishable an article is. The evolution of digital publishing and the medium of the Internet has not significantly changed the publication process in so far as it fundamentally works with the conventional written article; however, it has had a significant impact on how the academic community works with underlying data and in particular crystal structures.

This section considers the impact that the dramatic rise of the facility and the consequent increase in volume of data generation described in Sect. 1 has on the publication system, how we curate the data and how this can inform and influence the science that is being performed.

2.2 *(Crystal Structure) Data Becomes a First-Class Citizen in Publishing*

Until the early 1990s, crystallographic 'data' were generally included in articles as print-outs of coordinates, agreement factors, etc. and figures (with hardcopy structure factor tables submitted as supplementary information). However, from as early as the 1960s, crystallographic data was collated from the literature in a form of database to provide crystallographers with a more effective route to find structural information and a mechanism to bring the data out of the publication (or supplementary information). With digitisation came opportunities; however, crystal structures would have essentially been destined to obscurity of supplementary information like much of spectroscopic characterisation data today, if it had not been for the advent of the Crystallographic Information File (CIF) [56, 57]. Convened by the IUCr and driven by experts in the community, this standard for the representation of crystal structures was devised and essentially universally adopted [58]. CIF has found many uses in almost every aspect of crystallography-related research and is used by many different stakeholders. Through well-supported governance from the IUCr, CIF has moved way beyond just being implemented as a file format and become the Crystallographic Information Framework [59]. This innovation is exemplary and has resulted in crystallography being a leading light in data management and communication. Not only is it possible to describe an experiment and result with CIF, but it can support validation, visualisation, computation and curation. It has also been possible to write a paper in CIF for a number of years – this, alongside the other capabilities of CIF, has generated and led to a culture where it is entirely possible to 'publish' a dataset in its own right, without the need for a lengthy accompanying paper. The fact that the underlying data has risen to such independent prominence has led to IUCr convening a committee on data – CommDat [60].

The ability to automatically validate a small molecule crystal structure begins to provide some solutions to address the problems that having an increased number of structures in the system presents. The CheckCIF [61–63] service provided by IUCr is a key part of the community infrastructure that has been built on CIF. When

validation tools and services were first introduced around two decades ago, there was a tendency to use them as a binary gauge of quality, i.e. if a structure fails to meet all test criteria, then it is simply not acceptable. While this attitude can still be encountered, on the whole it has become standard practice to use this service as a tool to assist review and complement a reviewer's own judgement – that is, to use it in combination with other assessment approaches. These combined approaches to assess 'validity' of a crystal structure have become increasingly important as the fields of chemistry and structural chemistry have evolved over the last few years. Crystallographers are posed with increasingly difficult problems, where often it is not possible to recrystallise a product but crystallographic characterisation is crucial.

We are beginning to move to a culture, supported by the right processes and tools, where the notion that 'data needs to be fit for the purpose' is acceptable. In some areas of chemistry, any amount of structural information, e.g. connectivity or conformation, can provide huge insights and be immensely valuable, whereas in certain studies differences of a thousandth of an angstrom for a particular interaction can be very important. Clearly one needs to be able to make a judgement as to whether the evidence/data that supports a claim is of an appropriate level of quality to do so. We now have in place most of the standards, tools and processes to be able to do this. So long as the appropriate approaches are taken, it should now be perfectly acceptable to publish a structure that traditionally would have had an unacceptably high R-factor, so far as the structure is basically 'correct' and the author is not attempting to make grossly overstated claims in their analysis.

As a result, publishers are now only one aspect of the research lifecycle where quality is assessed. Database providers now perform more quality assessment before accepting deposits and as part of the process of validating structures for entry into their collections, but also end users now have the tools to deduce these quality levels themselves. These factors mean that now it should be possible to rapidly make crystallographic data available and if the right processes are followed and correct tools applied, then the end user can readily make use of this data with the confidence that it is being used appropriately. This is particularly important as we see the rise of computational/theoretical (a very significant proportion of the users of crystal structure data are performing in-silico calculations on them) and cheminformatics approaches in research – especially areas that are purely consumers of crystal structures.

However, we do need to be careful! If one considers that the crystal structure results we are talking about are in fact just models and based on how raw data has been interpreted, then there may well be cases where simply assessing and validating a regular result, i.e. a CIF, is not enough.

The small molecule crystallography community, through the IUCr CommDat, is now assessing in great detail the extent to which the raw data underpinning the CIF result needs also to be made available [64]. There are strong arguments and trends that this will happen as a routine matter of course in the macromolecular crystallography community; however, there are valid reasons why this might not be the right approach for small molecule crystallographers. Firstly, the sheer volume of results in small molecule crystallography is much greater, and so 'publishing' associated raw

data becomes unwieldy very rapidly – it does not scale well. Secondly, the nature of the small molecule result is very different to that of a macromolecular one – the resolution of protein structures is relatively low and there is often 'room for improvement', whereas the likelihood that relatively little extra can be obtained from a high-resolution small molecule structure is much higher. As methodology develops, in some cases very rapidly, the ability to reassess old data and get a better result becomes very real and very worthwhile. There are numerous situations where recourse to raw data could therefore be of very real benefit to future users of our data, many of which are listed below:

- When a result provides a contribution to chemical knowledge, but is poor quality
- In order to support a 'grand' claim, i.e. an unusual result that has not been previously observed or considered possible
- To support cases where modelling of disorder, twinning, incommensurate or modulated structures could be open to other interpretations
- To support analysis of the modelling of diffuse scattering
- To make available, e.g. disorder, twinning, incommensurate, modulated, diffuse scattering datasets so others/future generations can attempt to (re)solve them
- To provide supporting evidence for 'Advanced Experiments', e.g. charge density, high pressure, phase transition, gas environment and excited states
- When it is clear that future improvement, e.g. for many of the cases mentioned above, may be possible through developments in software and modelling
- To make available training sets and benchmarks for software and method developers

As access to fast networks and large volumes of data storage become ubiquitous and normal, many of the problems that have historically been perceived as blocks to routinely curating, making available and accessing raw data have gone. As mentioned previously that is not necessarily a case for making everything available but certainly is when to do so is clearly going to be of value.

Coupled with the recent technological developments described in Sect. 1, these digital developments have provided a platform that is fundamentally changing the application of structural information. Partly due to the fact that the technique is seen as a much more rapid and accessible method of characterisation by colleagues in synthesis and partly because specialists can achieve so much more, the number of articles containing crystallographic information has risen dramatically. The Crystallographic Information Framework is very much an enabler for this phenomenon, and the result is that being able to process and communicate at these new levels has led to entirely new and increasingly independent fields of study.

2.2.1 Structure-Driven Independent Research Fields

The speed, accessibility and capability of the technique have led to many more academic research groups being driven by crystallography and becoming almost wholly dependent on it. This has given rise to fields such as crystal engineering and

becomes the crucial underpinning of others such as MOF research, magnetism and supramolecular chemistry. Crystal engineering is now embedded as a mainstream chemistry topic, as evidenced by the long-term establishment of dedicated mainstream community journals such as the Royal Society of Chemistry's *CrystEngComm* [65] and the American Chemical Society's *Crystal Growth & Design* [66]. Furthermore, there are many more articles in high-impact general chemistry journals that fundamentally depend on chemical crystallography.

For such studies, it is generally necessary that one has to work with the 'raw' products of synthesis, i.e. recrystallisation or alternative methods are not viable, or study all the variants of a system/family, i.e. it is not possible to pick the best example. Furthermore, in such fields it is not uncommon to be working with extremely large and complex systems, often fraught with the difficulties of considerable quantities of solvent of crystallisation. These are just a handful of examples used to illustrate the point that beyond a decade ago, it would not have been possible to work with such systems using the crystallographic technology available at that time.

2.2.2 The Meteoric Rise of the Database

Databases are ubiquitous in crystallography [67]. The Cambridge Structural Database (CSD) has been collected, curated and provided by the Cambridge Crystallographic Data Centre for over 50 years and is the largest collection of single-crystal diffraction-derived results for chemical crystallography. Other collections exist which complement and in part overlap with the CSD including the Crystallography Open Database (COD) [68] and the Inorganic Crystal Structure Database (ICSD) [69]. Powder diffraction data and incommensurate crystal structures are collected by the International Centre for Diffraction Data (ICDD) and the Bilbao Incommensurate Structures Database (B-IncStrDB), respectively. The Protein Data Bank (PDB) for proteins is similar to the CSD in many ways, and in the early 2000s, it rapidly became a critical central tool in the field of bioinformatics.

In the beginning, the CSD was created by abstracting crystallographic data from publications. The establishment of agreements with publishers and the development of the CIF led to the archival and curation of the underlying datasets, but it was still a somewhat manual process. Technical developments and the establishment of automated workflows with publishers over the last decade have made deposition and linking to parent publications more digital/Internet compliant. The last 10 years have seen a dramatic increase in the number of structures in the CSD with the database doubling in size from half a million structures in 2009 to one million in 2019.

The role of the database has accordingly increased significantly over the decades. The derivation of knowledgebases has helped to provide a deep understanding of the solid state and allowed the generation of a set of principles and guidelines to aid the crystallographer. More than 10 years ago, these were the preserve of a few interested researchers. This work started with the derivation of 'standard values' for geometry of bonds and particular functional groups [70, 71] and developed into interaction

(Isostar) and geometric (Mogul) knowledgebases (vide infra). These tools are used in a number of ways such as in structure checking, conformation generation and refinement restraints. The significance of these tools has reached critical mass, and they have become mainstream applications embedded into a range of different research ecosystems. For example, the Hydrogen Bond Propensity tool and Full Interaction Maps are used by the pharmaceutical industry in their workflows for the evaluation of polymorphism risk.

2.3 The Effect on Crystallographic Practice

The recent advances in experimental equipment and approaches outlined above clearly have an impact on the body of structural data that the community can accrue. This phenomenon clearly has the potential to be completely transformative; however, it is not without associated difficulties, and there are several factors that the community need to address and change practice in order to account for the following aspects.

2.3.1 Rate of Data Generation

The speed at which diffractometers can collect data has ramped up, generally in a stepped manner coinciding with the introduction of revolutionary technology. The last decade has seen the largest step up of all, with data collection being at least an order of magnitude faster now than it was at the beginning. The facility that is built around these instruments must therefore adapt, and this is a very difficult process – coping with change is generally a problem, especially if that means adjusting or rebuilding infrastructure. So how does the modern small molecule crystallography facility cope with this increased rate of data collection and volume of results generated?

The macromolecular crystallography community has very successfully addressed this through increasing levels of automation, and nowadays, the approaches in this field are very different to those of 10 years ago. Looking to this field and making comparisons is therefore a valuable exercise – and much of the following section will be presented in this way. Some attempts at developing facility operations akin to MX approaches have been made, e.g. [47, 72]. However, the culture and purpose of much of the field of small molecule crystallography is very different and is not necessarily so suited to automation approaches. In the life sciences, crystallography is one of a set of tools that a structural biologist will use to address a very specific problem, whereas small molecule crystallographers look at a much wider-ranging set of problems and often operate as a service to those making new compounds and materials.

There are two significant impacts on the facility when the rate of data collection increases so markedly. Firstly, in-house data management approaches need to adapt

very rapidly. It is no longer viable for a facility to operate in a primarily 'manual' mode in terms of where data is stored, how it is moved around, how it is identified, how it is accessed, how it is monitored and how it is curated. Diffractometer control software has made some advances in these respects, but invariably the small molecule crystallography facility will have to develop mechanisms to do this. A range of solutions exist, for example, some basic scripting can ensure that data is sensibly and safely moved and there are numerous off-the-shelf automated mechanisms for storage, back-up and archival – often provided by the institution/organisation in which the facility is located.

However, the issues of identifying, and thereby being able to search and retrieve, data are often overlooked, and as the volume of data scales, this becomes a significant problem. A further aspect that becomes more intractable as the volume of data blooms is that of curation – how does one make sure data is migrated to the most recent formats and storage media; how does one keep track of what data 'belongs' to whom and what you are 'allowed' to do with it; how do you know what has been published and what has not? These matters are not easily solved by a piece of technology and largely still have to be addressed by knowledgeable staff.

These factors also lead into the second main impact – that of the publication bottleneck. For some time now, the pace at which structures are generated has been greater than that at which they can be published. Despite the increase in digitalisation of much of the publication process, there has barely been increase in speed of publication, and the process has essentially remained unchanged for decades – particularly in respect of the 'publication' of data. In chemistry and small molecule crystallography, there persists a culture where a crystal structure is tied to the publication of the associated chemistry, and often the latter takes a considerably longer time to perform and understand. Accordingly, data generation far outstrips data publication, and facilities are building up vast archives of redundant data. There have been some steps towards addressing this situation, but there is still a long way to go in terms of cultural change. With the advent of the World Wide Web, researchers have experimented with self-publishing, from putting some results on a personal website through to depositing significant volumes in online repositories. The rise of general repository systems such as Figshare [73] has had quite an impact in some fields, although in general these generic systems tend to cater better for academic institutions, publishers and similar large organisations. However, the advantage of using a domain-specific repository, where available, is that data can be validated, checked and interpreted appropriately, so while in small molecule crystallography there have been some local- and project-level attempts to develop repositories, e.g. eCrystals and related projects [74], they have largely failed to gain widespread traction. However, CCDC is now beginning to address this, with a marked uptake in 'CSD Communications' [75]. For years there had been a 'private communication' route to deposit structures in the CSD that did not have an associated literature publication, but this was largely not used. With the cultural shift we have seen recently in terms of availability of digital resources, CSD Communications, due in part to the provision of an enhanced deposition process, has begun to thrive and now numbers around 30,000 entries. The ability to independently make a

crystal structure widely available through a formal and accepted route clearly creates some powerful opportunities to address a number of the problems we currently face.

Despite increasing digitisation and automation, in small molecule crystallography, it is unavoidable that an increase in the rate of data production requires input from larger numbers of skilled people. The requirement for skilled intervention and oversight in data management has been mentioned above, but there are other crucial aspects that currently need more 'pairs of hands' to address the increase in data collection rate. Many facilities need to optimise their operational time to be justifiable, and slightly perversely rapid data collections can be a problem here! Selection of crystals remains a key, rate-determining stage, and while more effective and automated screening of samples can now occur on the diffractometer, still finding and mounting the most appropriate crystal can be laborious and technically quite difficult, requiring numerous iterations and a lot of expert decision-making. Macromolecular crystallography tends to take the approach of measuring everything because that can be done rapidly and then only using the best data. The diverse nature of chemical samples makes this approach much more of a challenge to automate; however, there are still aspects of the philosophy of it that could be adopted in chemical crystallography.

An obvious way to optimise the efficiency of the modern crystallographic facility is to have many more skilled users – it might be argued that the era of the 'lone service crystallographer' is rapidly disappearing. In some cases, a large research group will have its own instrument and everybody will use it – in much the way macromolecular crystallography has run over the years. However, for the service crystallography facility, this is not the case, but there is an increasing trend for users of the service being sufficiently trained in order that they can perform the data collection themselves. Here the role of the expert crystallographer is changing to be that of an educator as well. However, while it is relatively easy to train a novice how to operate a diffractometer, this by no means helps with the task of processing and analysing the arising data. There are such a wide variety of difficult problems (sometimes generated by collecting poor data) in structure refinement, many of which are very difficult to generalise, that mean for the foreseeable future this aspect of the process will be where the main bottleneck arising from increased rate of data collection lies.

2.3.2 Quality of Data Generated

In the modern facility, the quality of data generated has a much greater spread than historically. In past eras when instrument time was rare and precious and had its limitations, there was a strong tendency to collect data on the better-/best-quality samples only. Modern instrumentation removes much of the instrument time availability problem and in turn allows one to address more challenging problems. Furthermore, there is a strong demand in certain currently fashionable areas of chemistry for any kind of result to add to the body of evidence to support claims. Also, the way in which the technique supports different disciplines has an effect on

data quality. The chemical crystallographer will often 'get what they are given' with limited opportunity to grow better crystals, whereas the macromolecular crystallographer will robotically attempt to grow crystals of the same sample in high-throughput modes under as many different conditions as possible.

These factors naturally lead to a much wider range, in terms of the quality, of crystal structures produced by modern facilities. However, the ability to validate results to a greater extent has also progressed. Due to the increased standardisation and digitisation of data and results, there are now not only a larger number of quantifiable metrics to assess quality, but also they can be combined and compared to further probe and measure quality. A couple of decades ago, a structure would be judged by just a handful of metrics, and often the non-expert would only consider the R-factor. However, modern validation tools are now accessible to all and results presented in such a way that it is relatively easy to judge quality based on a whole range of factors. So, one might now consider that it would be possible to publish a much wider range of quality of results, if not all that can be deemed as 'correct'. This is a crucial change in the mindset of publishing, but a key one to support modern chemistry requirements.

This gives rise to the concept that 'data need to be fit for purpose'. The primary purpose of a service crystallography structure in the past was essentially a proof of what has been synthesised. Increasingly, over the years, crystallography has moved on from this configuration/conformation requirement to look in detail at the nature of bonding and crystal structures in the form of intermolecular interactions and thus requiring a certain quality of result. So, it should be perfectly acceptable, for example, to publish a structure that critically reveals the stereochemistry of a compound but does not reliably provide any detailed insight into bonding. Of course, the structure must be correct, but it can be of low quality – and these factors can readily be assessed.

Furthermore, in the modern era, it is now critically important to consider that a result has a very (in some cases, more?) important role when incorporated in and contributing to the whole body of knowledge, e.g. a structure becoming a component of a database. There are illustrations throughout the literature, and this article, of the benefits of being able to analyse a database to derive rules and provide the basis for informatics. In an ideal world therefore, every structure determined should in principle contribute to this body of knowledge.

We can make a big step towards this goal. In the modern world of data science, it is very common to spend a significant amount of time 'cleaning' and 'processing' a large dataset. If we are to move towards a 'data-fit-for-purpose' approach, then validation becomes an imperative process. This can readily be achieved not only using our modern validation tools but also by making all of our data available. In recent years common practice has shifted, over a relatively short period of time, from provision of structure factors being very rare to this being close to mandatory for most publishing routes. Deposition of structure factors has empowered a whole new level of validation, where not only the model can be compared to the data from which it was derived but also refinements can be performed, which opens up a further level of validation. This has moved from a culture of 'having to take

someone's word for it' to 'caveat emptor', i.e. it is possible to take any data and understand it enough to be able to trust it, or not. 'Poorer' structures therefore need to be supported by underlying data, and significant advances have been made with structure factor deposition, but in certain cases underpinning with the raw data would also be desirable. Given that a set of structure factors are extracted from diffraction images, the user of a result still has to trust that the generator of it has done the best possible job. In some cases, it may be possible to reprocess using different routines or approaches and extract better structure factors. For particular types of problem, e.g. where diffuse scattering is prominent, this would be highly beneficial. Indeed, in many such cases, it would also enable the future potential to do better by supporting and driving software and algorithm advances.

2.3.3 FAIR and Open Data

There is a moral requirement to make the outputs of funded research easily available for the purpose of sharing and re-use. Recently there has been significant growth in a strong global movement across all disciplines to make all research data/outputs FAIR [76], that is, Findable, Accessible, Interoperable and Reusable. The FAIR principles [77] are intended as a guide and detail for each of these aspects the key factors that need to be considered and in place to realise them. Specifically, the 4 FAIR foundational principles can be explicitly and measurably described by 15 FAIR guiding principles, and any interpretation or implementation of these principles may be chosen, so long as they lead to machine-actionable results. It is important to note that these principles can readily be misinterpreted if they are not embraced with the spirit by which they have been derived, e.g. FAIR is not a standard, FAIR is not just about humans being able to find data, and FAIR is not equal to Open [78]. Some examples from the 15 guiding principles that are relevant to service crystallography and which the crystallographic community has been working towards for some time include:

F1 (meta)data are assigned a globally unique and persistent identifier;
F2 data are described with rich metadata;
F4 (meta)data are registered or indexed in a searchable resource;
A2 metadata are accessible, even when the data are no longer available;
I1 (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation;
I2 (meta)data use vocabularies that follow FAIR principles;
R1.3 (meta)data meet domain-relevant community standards

At the discipline level, chemistry is not well prepared to implement these principles [79]; however, crystallography is very well established in this respect. For some time, the notion of 'good data stewardship' has been part of chemical crystallography practice. The management of data in-house, that is, within one's own laboratory, is generally ad hoc but with some implicit principles, but at centralised facilities and in some respects in the public domain, the community has established data stewardship practices which are considered exemplary.

One might argue the many aspects of crystallography have been FAIR since the acceptance and widespread adoption of the CIF format, which clearly addressed the machine-actionability aspect of FAIR well ahead of its time. It is only recently that raw data have been considered as FAIR and the community is avidly in discussion about why and when one might make raw diffraction data openly and FAIRly available. However, while CIF addresses aspects of interoperability and reusability and technically enables data to be findable, it is worth noting that it is only recently that the community has begun to consider the *Open* aspects of FAIR. While our data is well structured and tools exist to explore it very well, they have not been essentially open. This is illustrated by some of the other FAIR guiding principles:

F3 metadata clearly and explicitly include the identifier of the data it describes;
A1 (meta)data are retrievable by their identifier using a standardized communications protocol;
A1.1 the (communications) protocol is open, free and universally implementable
I3 (meta)data include qualified references to other (meta)data
R1.1 (meta)data are released with a clear and accessible data usage license
R1.2 (meta)data are associated with detailed provenance

Social and financial attitudes are beginning to change; however, there is still a considerable amount of work to be done. A truly FAIR approach is difficult to achieve, particularly in the home laboratory where often a single crystallographer is operating under constrained financial and political circumstances. However, funders of research are now beginning to realise this and are in fact driving the FAIR agenda. Many funding agencies, e.g. NSF [80] and UKRI [81], with the European Research Council [82] generally being the most forthright, now provide support, admittedly generally within large grants, to ensure FAIR data results from their funded research. This funding and its associated mandates ensure not only that individual laboratories are able to devote appropriate resources to good data stewardship but also that via a full economic costing model it enables institutions to support its researchers to work in the spirit of the FAIR principles. Crystallographic databases are becoming increasingly aware of the importance of the FAIR principles and good data stewardship and are currently rapidly developing their strategies in this respect [83].

2.3.4 Complexity and Diversity of Chemistry

Finally, given that service crystallography is an underpinning technique, it is worth considering in brief the extent to which the field of chemistry has developed in recent years. There is an ever-increasing drive from funding agencies and society for fundamental research to address real-world problems – this has led to an increased complexity in the nature of chemistry research being undertaken. Furthermore, this drives a greater need to understand the application of materials in 'real-world' situations. These drivers produce significant challenges for crystallographers – both in the sheer size and complexity of the structures being generated and often in their dynamic nature. This often means molecular structures are much larger and can be comprised of multiple components – meaning they can be flexible and adopt

numerous conformations. The rapid rise of MOFs (vide infra – this class of material now represents about 10% of the CSD entries; see Fig. 6) provides another example of highly functional materials that pose challenges for the crystallographer – such as flexibility, large void spaces and significant amounts/proportions of solvent. Porous materials, single molecule magnets, frameworks and supramolecular complexes are all examples of areas of chemistry that have grown fantastically fast over the last two decades and that rely fundamentally on crystallographic analysis.

The advances in instrumentation outlined above provide an excellent basis for addressing these challenges. The strength of X-ray sources, combined with noiseless, large dynamic range detectors, can produce a far clearer and greater quality standard of data at a much faster rate. In some senses, this empowers the crystallographer; however, in many cases, this produces even greater challenges when attempting to analyse this data. Having instrumentation that can make more detailed measurements on more challenging systems is now tending to expose either the complexities of these materials or the complexities of the solid state itself. This means that the number of cases of disorder is rapidly rising and that they are becoming more complicated (as discussed later in Figs. 3 and 8). Furthermore, the incidence of modulated and incommensurate structures is increasing and becoming more apparent in molecular systems – this is a phenomenon that has been considered very rare until recently and also one that we are poorly prepared to address.

2.4 Future Considerations

These shifts in instrumentation and chemistry being undertaken clearly indicate two needs for future development to focus on. While software has been well developed to drive diffractometers, mainly by manufacturers of this instrumentation, there is much to do regarding processing and working up data. This is generally the part of the process where the most time and expert input is required and often the stage when a study is aborted. In many of these cases, there would be potential to extract some structural information albeit possibly of a different standard to that which is currently demanded.

Firstly, there is a shortfall in the drive to technically develop data processing algorithms and approaches. The advent of new detector technology highlights the fact that many laboratories are using integration and correction routines that were designed for CCD data over a decade ago. Currently this is appropriate for the majority of use, but for the HPC detectors that will be commonplace in the future, we are essentially using the first generation of this software, and there are developments that can be made. There are also different approaches that could be used or developed in order to get a lower level of structural information from particularly ‘poor’ diffraction patterns. Furthermore, as photon counting and next-generation integrating detectors gain more widespread use, it will also be necessary to devote more attention to data processing methods.

Secondly, structure solution routines are now quite powerful and can generally provide an appropriate starting point for structure refinement – which in a large number of cases cannot be satisfactorily completed. There is therefore a requirement to devote effort to improving and developing approaches to modelling and refining such structures. Structure refinement software has for many years been driven by the few and taken for granted by the many. The community now faces a grave problem in that there is dwindling expertise in the area of crystal structure refinement software development. This is in part due to the 'golden age' of crystallographers who revolutionised service crystallography and solid-state structure characterisation in the 1960s, 1970s, and 1980s being at the end of their careers. There has been little or no business or academic reward incentives driving software development in the intervening decades, resulting in a degree of stagnation. There have been some advances in structure refinement, e.g. rigid body and new constraints or restraints, with Olex2 [84] and CRYSTALS [85] introducing some notable tools, but much more development is required and particularly for the newer challenges that single diffraction is generating. In the modern digital era, the need for software developers has been recognised in society and in academia to an extent – and our subject area is no exception.

3 The Database Revolution

3.1 Nature of the CSD

Several databases of crystallographic data exist, including the three detailed here. The Cambridge Structural Database (CSD) [14] contains the world's collection of organic and metal-organic small molecule crystal structures collected and shared by the CCDC (Cambridge Crystallographic Data Centre). Inorganic structures are collected in the Inorganic Crystal Structure Database (ICSD) [69] provided by FIZ Karlsruhe. Protein structures are contained in the Protein Data Bank (PDB) [86]. As the majority of chemical crystallographers concentrate primarily on molecular materials, this article concentrates for the most part on the CSD.

The CSD was founded in Dr. Olga Kennard's group at the University of Cambridge in 1965 following inspiration from JD Bernal that the 'collective use of data would lead to the discovery of new knowledge; transcending the results of individual experiments' [87]. This has been seen in the growth of the CSD, with the database doubling in size from half a million entries in 2009 to one million in 2019. The cumulative growth of the CSD can be seen in the stacked columns illustrated in Fig. 2. An analysis of the percentage of organic and metal-organic structures shows that the number of metal-organic structures (38,131) first overtook the number of purely organic structures (37,733) in the database in 1989. Although the ratio of organic to metal-organic structures across the whole CSD has not changed greatly since 2009, metal-organic structures currently account for 56.9% of the CSD.

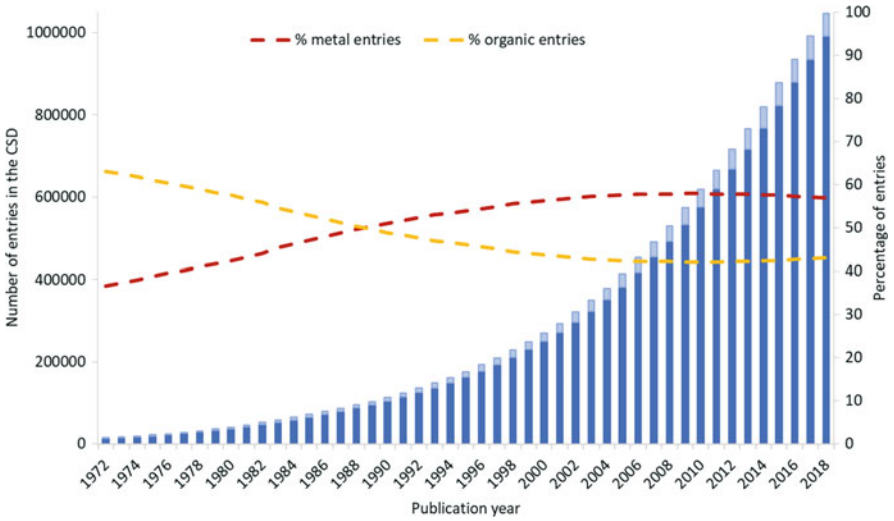


Fig. 2 Number of entries in the CSD and entries added to the CSD (dark-blue column and light-blue column). Overlaid the percentage of the CSD that is organic (yellow dashed line) or metal-organic (red dashed line)

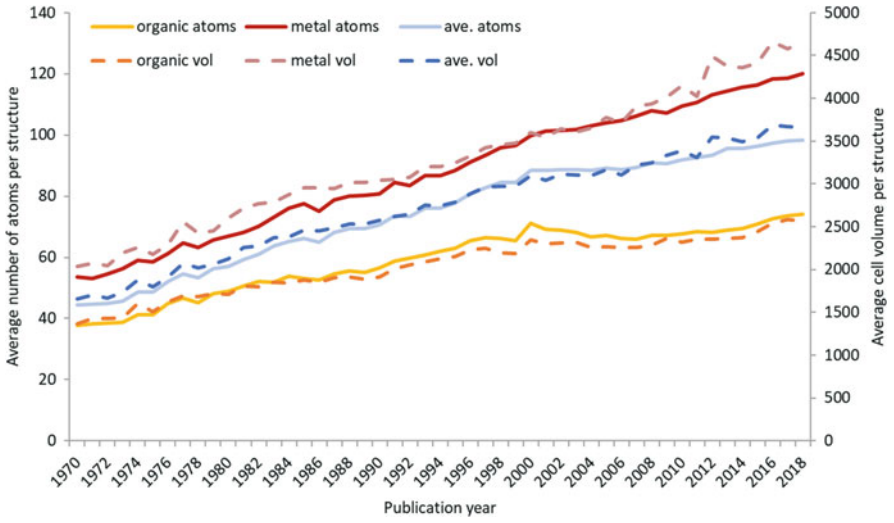


Fig. 3 Trends in the average number of atoms per structure and average unit cell volume (Å^3)

As alluded to in Sect. 2, it is evidenced in the CSD that it is not just the number of crystal structures that is increasing but the size, complexity and diversity of structures are also growing. One way to demonstrate this is by looking at the average number of atoms per structure which shows an increase in the size of the compounds determined; see Fig. 3. With the increase in the average number of atoms in a

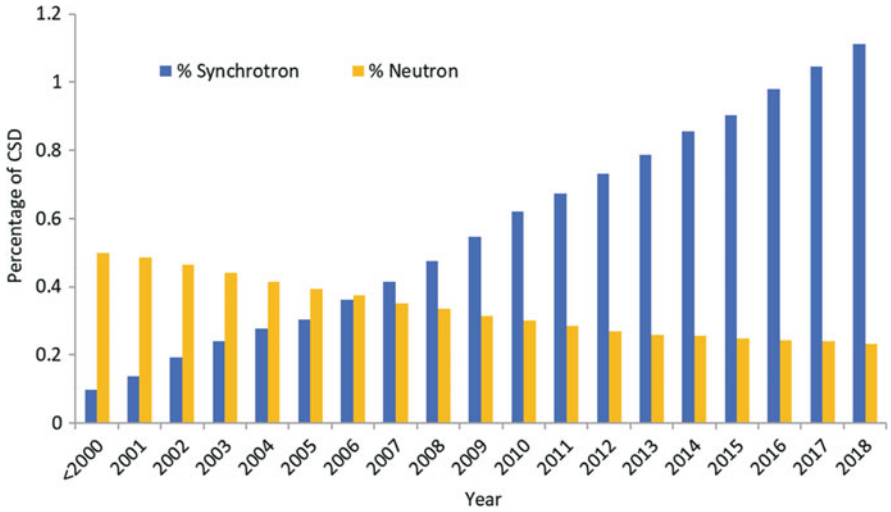


Fig. 4 Trends in the percentage of synchrotron (blue) and neutron (yellow) structures added to the CSD per year

structure, there is a matching increase in the average unit cell volume. It is clear that in recent years this increase is being driven by metal-organic structures, as the average number of atoms per structure for organic structures shows little change since the beginning of the twenty-first century. Interestingly, the average density has not increased and has been steady at an average of 1.55 g/cm^3, varying only by 0.02 in the last 10 years.

The ability to solve many of these bigger structures can largely be attributed to the improvements in modern crystallographic software, methods and instrumentation as outlined in the earlier sections. Central facilities are playing an increasing role in the structures that are added to the CSD, as can be seen from the growth of synchrotron structures in Fig. 4. While the percentage of the whole CSD from neutron studies is decreasing, the number of structures added each year is increasing, particularly in the last 5 years. The CSD also contains a small number of electron diffraction studies, and this number may increase as the technique develops in the future. Current work is underway to investigate the output from individual synchrotron facilities and will be discussed in a separate publication. This work has however highlighted that the current mechanisms for connecting information concerning the recording of data collection at central facilities and its subsequent publication could be improved and that the provenance of the facility used is not well captured as part of the crystallographic result.

As the range of structures solved by crystallography has diversified and evolved, so too has the range of space groups in which they crystallise. Over half of the structures in the CSD are monoclinic, which is due to the high numbers of structures in $P2_1/c$, as well as C2/c and $P2_1$ (see Fig. 5). Over a quarter of structures are solved in triclinic space groups, predominantly from P-1. A growth in P-1 structures has

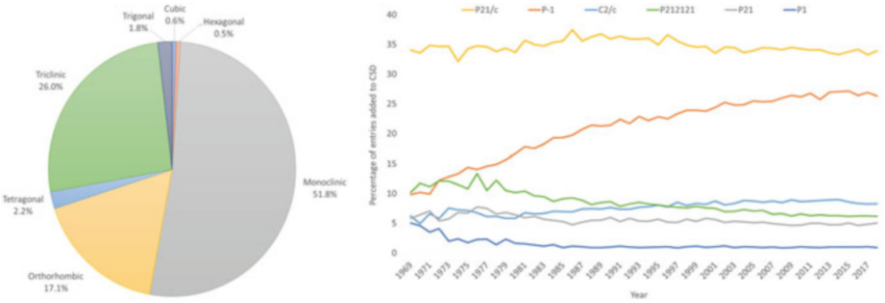


Fig. 5 Composition of crystal classes and changes in the frequency of some of the most common space groups in the CSD over time

been seen despite papers which have looked at missed symmetry [88–90] and has led to new tools to aid in the identification of crystal symmetry [61]. It is notable that overall the gap between the percentage of chiral and non-chiral space groups is slowly increasing and is now grown to 83.7% non-chiral versus 16.3% chiral. The number of traditionally natural chiral compounds, such as amino acids and steroids, added to the CSD per year has remained fairly constant, while there has been a continual increase in the number of structures added to the CSD each year.

3.1.1 Chemical Composition

This section considers what the instrument development has meant for the chemical makeup of the structures in the CSD. The elemental composition of the CSD is heavily weighted towards lighter non-metals commonly found in organic compounds. The percentage of structures containing lanthanides and actinides has seen a large increase since 2009, reflecting the use of these elements in materials prominently used in research into new technologies such as hybrid cars, superconductors and single molecule magnets. However, the relative proportion of lanthanide- and actinide-containing compounds in the CSD overall is still very small (Fig. 6). In the last 10 years, there have been structures published that contain four new elements not previously observed in the CSD. These are helium, added in 2013 (CSD-YEMTUH) [91], berkelium added in 2016 (CSD-EVAMIZ AND EVAQUP) [92], californium first added in 2010 (PUTQEB [93], HIYLAE in [94], IHAJOS in [95] and FIHLIU in [96]) and neon which was first added in 2016 and now appears in a total of 22 structures in the CSD.

Delving a little bit deeper into what types of compound are most represented in the CSD, Fig. 7, indicates changes in the number of structures of common types of compound. There is a big increase in the number of MOF-type structures since 2000, with a jump from 2.5% to 9% of all structures in the CSD included in a subset of MOF-like structures [97]. This reflects the versatility and ease of synthesis of this class of materials, with increasing interest in their application in sustainable

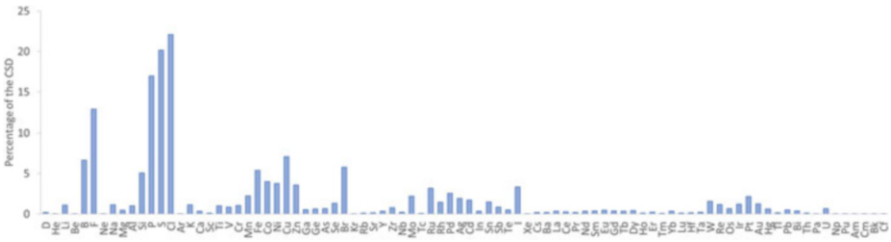


Fig. 6 The percentage of structures in the CSD that contain each element, excluding C, H, N and O

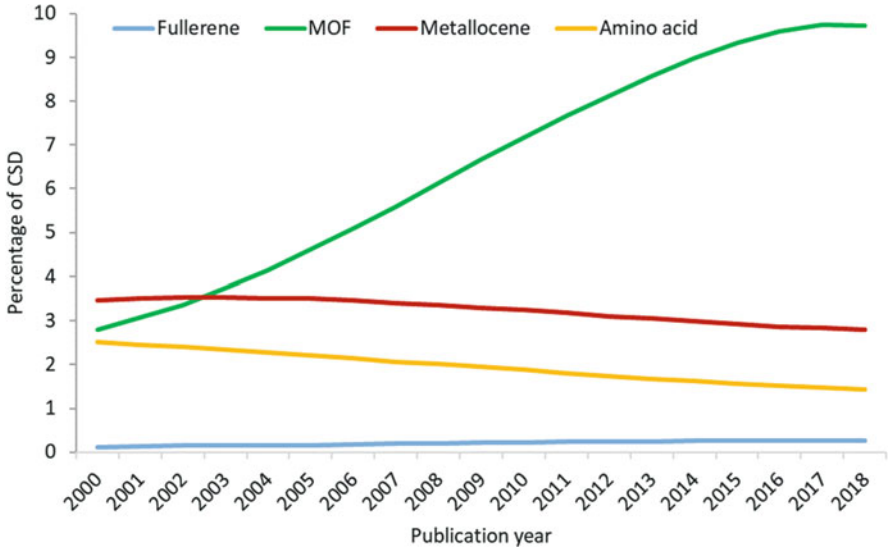


Fig. 7 Percentage of structures in the CSD containing common classes of materials

technologies such as gas storage, purification and separation as well as in catalysis and sensing [98]. It also reflects the fact that often the only comprehensive form of characterisation for MOFs is by X-ray diffraction in the solid state as NMR techniques to not give full structural information and so a crystal structure is often a crucial component enabling publication.

3.1.2 Effect on Model Refinement

It might be expected that, with an increase in complexity, the agreement between the model and the experimental data would also follow the same pattern. However, the average R-factor (R1 is discussed in this section) has not changed much over the last 10 years with an average of 5.28% up until 2009 compared to 5.12% from 2009 to

present. It is possible that higher R-factors observed for more challenging structures will be offset by the fact that there will be an increasing number of very good values resulting from advances in instrumentation, technique and software and therefore the net R-factor remains approximately the same. However, the consistent value at around 5% may also be defined by a general requirement from some journals for structures to be of a certain, relatively high, standard and that the community is not embracing the 'data-fit-for-purpose' concept. It is acknowledged that the R-factor is just one indicator of data quality and others should be taken into account when the data is considered for re-use. Of course, the chemical sense of the final model should not be overlooked, via a consideration of atom assignment, analysis of molecular geometries and an appraisal of anisotropic displacement parameters (ADPs). Using a range of metrics would enable an overall assessment of the final structure quality over the whole process from collection to refinement [99].

The increase in the size of molecules, together with improvements in technique and software, has resulted in an increase in structures with modelled disorder. If the current trend continues, it is estimated that 50% of new structures added to the CSD will have modelled disorder by 2060. Modelled disorder appears to be more common in metal-organic structures, as 70% of disordered structures contain a metal. Highly symmetric components and those without strong intermolecular interactions to neighbouring molecules may have greater freedom to occupy multiple orientations throughout the crystal and therefore exhibit disorder which is difficult to model. Perhaps due to the rise in popularity of MOFs, there has been a significant increase in the number of structures that contain a co-crystallised solvent, or solvents, in the structure. These too are invariably very disordered, often adopting many orientations in large void spaces in this type of structure. It is now possible to omit such entities from the calculation of structure factors and hence from the model itself. Since A.L. Spek's initial paper [100] on the BYPASS routine and the subsequent implementation in PLATON [62] and Olex2 [84], there has been a dramatic uptake in the use of these approaches. This is illustrated by the increase in the percentage of entries using the SQUEEZE or MASK algorithm in refinement to remove solvent and produce a better model fit for the part of the structure of primary interest (Fig. 8). This has allowed complex structures to be determined that previously would not have been possible, such as those with severe disorder and those where a satisfactory model for some of the electron density cannot be found.

One effect that the use of improved apparatus as described in previous sections of this article can have is in the observation of improved precision for bond lengths. This could be illustrated by the increased percentage of structures with low average estimated standard deviation on C-C bonds. Increased precision is valuable for a variety of reasons; however, it has a significant effect on the re-use and repurposing of data in building more accurate models and knowledgebases, which are amongst concepts that will be discussed later in this review. The inclusion of hydrogen atoms in the refinement model is also valuable for the re-use of data, especially when examining hydrogen-bonding motifs. Improvements in data quality and refinement procedures have resulted in a dramatic increase in the inclusion of hydrogen atoms

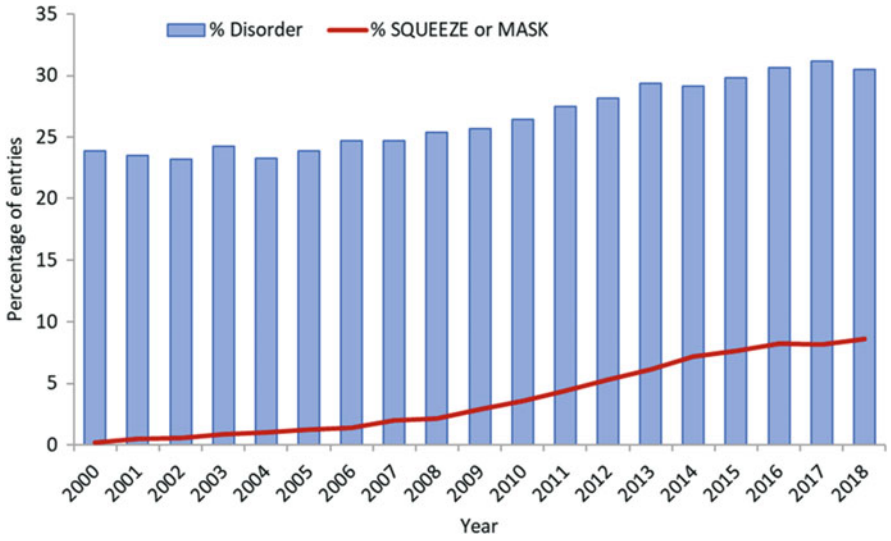


Fig. 8 Trends in the percentage of structures in the CSD that are disordered and percentage of structures added that have used the SQUEEZE or MASK routine (red line)

since the CSD began – over 90% of structures in the CSD now have hydrogen atom coordinates included.

3.2 The Evolution of the Structural Database

It took 5 years to collate the first volumes of the CSD which were published as a book series, *Molecular Structures and Dimensions*, in 1970. It was soon realised that to make the most of the database, it would be useful to be able to search and retrieve collected structural data electronically, and a computer system was developed [101]. This section highlights the current state of the art, alongside some key developments in the intervening years, in key areas where an electronic database system adds value to the collection of data.

3.2.1 Visualisation and Analysis

Crystallography is a powerful technique, providing accurate and unambiguous information about the 3D nature of molecules and solid-state crystal structures. For this reason, visualisation of structures is very important. A large number of visualisation software packages have been made available over the years; some examples include ORTEP [102], Diamond [103], CrystalMaker [104], XP [105],

CAMERON [106], ORTEPIII [107], X-Seed [108] and Olex^2 [84]. CCDC recognised the importance of crystal structure visualisation with the release of the program Pluto [109] which was superseded in 2001 by the current visualisation and analysis program Mercury [110, 111]. Mercury enables the visual exploration of crystal structures and includes features such as the ability to generate packing diagrams, build and explore networks of intermolecular contacts, display space group symmetry elements and calculate and display voids and, since 2019, polyhedral display of metal-organic coordination compounds. Initially, any analysis of the data had to be performed in external applications and then imported into a visualiser; however, the launch of Vista in 1994 incorporated both these elements [112]. Vista was later incorporated into Mercury as the data analysis module [113]. Further developments to Mercury [114] included the ability to display ADPs, calculate powder patterns, overlay two structures for the purpose of comparison and display predicted crystal morphologies.

3.2.2 Database Searching

Initially the *Molecular Structures and Dimensions* book series provided a way for chemists to look up structures based on bibliographic details and compound names, much like an encyclopaedia. This allowed crystallographers to check which structures had previously been determined. In 1988 QUEST [115] was released and provided integrated search facilities for text, numeric and 2D chemical information. Notably it enabled chemical similarity and substructure searching for the first time from a 2D sketcher. The graphical user interface was modernised in 2002 to the program that is still in use today, ConQuest [110]. The power and versatility of substructure searching has made it an invaluable research tool that has been a core part of the CSD system ever since, and indeed the idea of not being able to search all structures containing a drawn fragment would be challenging for many modern research purposes. Today similarity searching is a common technique that is available in a number of programs, and there are several methods and coefficient scoring measures that can be employed [116].

Information from the CSD can be used to inform the direction of new experiments. CellCheckCSD [117] is a command line tool for checking unit cells against the known structures in the CSD during data collection, which allows a user to determine if the compound is novel, accidental crystallisation of starting materials or a reaction by-product. This has been integrated for automated for use through the Rigaku software package CrysAlisPro and Bruker software suite APEX2, and its utility in the experimental screening process is discussed in Sect. 1 [118].

The complexity of searches that may be performed has undergone a large shift in the twenty-first century from the early simple ‘look-up’ searches to more advanced searches which enable scientists to gain new insights from the data. One of the first leaps was the addition of Motif searching [119] in 2004, in which a 2D substructure search is followed by a 3D geometric search to find specific intermolecular interaction patterns. This capability was a significant tool that fuelled the growth of crystal

engineering and spurred activity investigating molecular packing in the solid state. Packing similarity [120] became an important factor to assess whether different crystal structures could be polymorphic. The XPac [121] approach went further to assess the degree to which crystal structures formed by different compounds might be closely related. Implemented in Mercury around the same time, COMPACK [122] looks at the similarity over multiple molecules when comparing crystal packing. The related packing feature search [120] in Mercury will retrieve structures with a similar crystal packing feature to that selected, such as a hydrogen-bonded dimer. Other approaches of comparing crystal packing have developed, CrystalCMP [123] employs differences in molecular rotation and relative positions, and Salbego et al. proposed an index for 'supramolecular clusters' [124].

As computers have developed, the types of analysis scientists would like to do have become more complex. There has been a trend towards informatics with the rise of 'big data' and using computer programming to directly search and analyse the database. The CSD Python API (Application Programming Interface) [14] has enabled this type of query; with much of the CSD functionality now available through fast computer algorithms, the knowledge has become easier to discover. This means that it is possible to use most of the database searching functionality in a machine-to-machine mode, allowing for scripting and automated workflow enactment and providing the potential to power machine learning and big data approaches.

3.2.3 Web Tools

Technological advances and the proliferation of the Internet led to online access to the database through WebCSD in 2010 [125], which gave easy access to the CSD via a website without needing to install software locally on the user's computer. This has been superseded by a second version that reflects the changes in web technologies. The ICSD [126], PDB [127] and COD (Crystallography Open Database [68]) [128] databases provide a similar level of text and bibliographic searching via a web interface. The distribution of the CSD and associated tools has also been transformed by web developments, which have been vital as it has grown in size – in 2009 the CSD was distributed on a CD, whereas today it is installed by download from the Internet.

CrystalWorks [129] is a web-based search tool that takes advantage of the fact that the different chemical crystallography data sources often need to be searched in the same way. By indexing and aggregating the CSD, ICSD and COD, it enables a simultaneous metadata search across them all at a high level, e.g. bibliographic, composition or unit cell queries. It should be noted that the full CrystalWorks is currently only available to UK academics through the Physical Sciences Datascience Service [130]; however, there is now a developing interface for all that provides access to open (small molecule) crystallographic data. Conversely, while there are a lot of similarities in how different crystallographic databases are searched, there are also disciplinary differences, for instance, in the cases of the ICSD, the

PDB, ligands and sequences and the CSD molecular, packing and interaction similarity. The Internet has also become a vehicle for educational resources based on crystallographic data, e.g. of particular note are the symmetry resources produced by Professor Dean Johnston at Otterbein University [131]; the International Union of Crystallography (IUCr) managed list of numerous online educational resources [132]; other society resources such as the BCA [133] and ECA [134]; and resources linked to the crystallographic databases [135–137].

As mentioned previously, a key development that aided the storage and analysis of crystallographic data was the innovation of the Crystallographic Information File (CIF) [56, 57]. This standardised computer and human readable format was implemented by the IUCr for their journals and rapidly taken up by the community. Less than 10 years after its launch, over 90% of the structures added to the CSD in 2000 were deposited in CIF format. Newly refined data can now be emailed directly from Olex2 to request CCDC numbers ready for inclusion in a publication manuscript. The use of CIF has enabled automation of some processes by computers including a web deposition service [138], introduced in 2010, for sending the finalised file to CCDC. This was an improvement on the mailing of paper copies previously used and has enabled automated checking of the data as it is deposited. Following further enhancements, the service is now fully interactive providing feedback to users through the process and is designed to help increase the quality of data and incentivise depositions. Today the service links up to ORCID so users can associate their researcher ID to their datasets; it checks the syntax of submitted files and allows the author to add extra information including a link to raw diffraction data. The eight-step process also enables the researcher to check the unit cell against existing structures, add crystallographer details to ensure the crystallographer and send the file through the IUCr's CheckCIF [61, 139] service to check the integrity of the data.

3.3 The Transition to Informatics: Methods and Tools to Leverage Databases

3.3.1 Knowledgebases

Information on individual crystal structures is useful, although much more can be learnt from the trends revealed by combining the data together. Extracting this information and compiling it into a form that can be used to compare a given structure to many others that have similar chemistry leads to the formation of derived libraries of 'knowledge'. The CSD has collated two knowledgebases, for molecular geometry and intermolecular interactions.

3.3.2 Molecular Geometry

As structural data was aggregated, tables of derived data could be formed. In the late 1980s, there was an enormous effort by Orpen et al. that produced tables of precise bond lengths for organic [70] and organometallic [71] compounds. These were updated in 2006 in the International Tables for Crystallography [140, 141]. The use of these tables to check geometry in newly determined structures was a laborious process and still involved manually looking up the data for individual bonds. Leading on from this in 2004, a derived knowledgebase of geometric parameters was developed by the CCDC called Mogul [142], which facilitated the automatic retrieval of bond lengths, angles and torsions and was further enhanced by the addition of data on ring conformations [143]. Mogul is now widely used in checking new small molecule structures for unusual geometry by comparison with distributions of known geometry for similar fragments derived from data in the CSD. A similar process is also embedded in the PDB deposition pipeline to check the geometry of ligands in protein structures. Utilising this data can be taken further in the application of a knowledge-based conformer generator [144, 145] which allows the minimisation of molecular conformations and the generation of conformer subsets based on CSD data.

3.3.3 Non-bonded Intermolecular Interactions

The first computerised library of intermolecular non-bonded interactions was collated in Isostar [146]. This reveals the geometric preferences of hydrogen bonds and other classical directional non-bonded interactions. Since 2009 further interactions have been added, such as C-I groups, due to a significant increase in interest in non-classical intermolecular interactions, such as sigma-hole interactions [147–151]. It is important that a knowledgebase can potentially grow and the value of a greatly increased quantity of crystal structure data is demonstrated in Fig. 9, where it is clear that more data points provide a considerably more conclusive understanding of preferred directionality.

The most significant evolution from Isostar was into Full Interaction Maps [152], which allow the concurrent visualisation of multiple different non-covalent interactions. Full Interaction Maps use Isostar data to construct contour density maps around a whole molecule, as seen in Fig. 10. These maps can be used to evaluate the stability of polymorphic structures, assess multiple types of non-covalent interactions simultaneously and provide a platform for understanding crystal morphology. A comparison of how well the hydrogen bonds of a given crystal form coincide with the likely interactions highlighted by Full Interaction Maps can indicate the stability of the form. This approach is now used by pharmaceutical companies and the CCDC to help determine the stability of a new drug form and the risk of a more stable polymorph being found with different physical properties [153, 154].

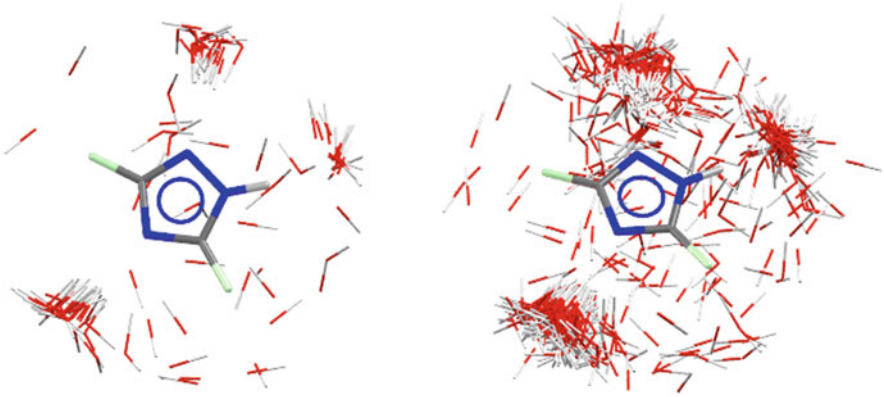


Fig. 9 Isostar scatter plots for triazole with any OH contact group in 2009 left and 2018 right

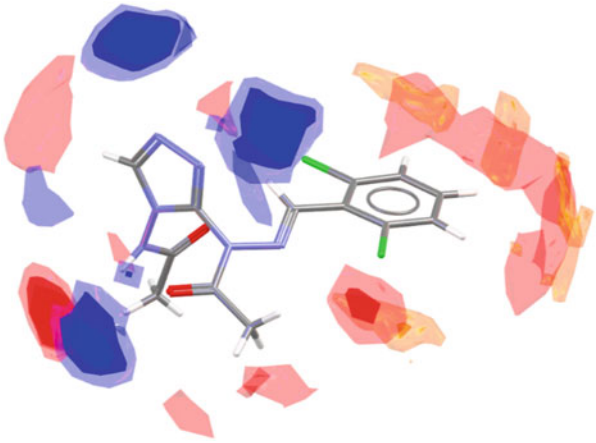


Fig. 10 Full interaction map for AABHTZ showing areas where interactions are frequently seen for uncharged N-H (blue), carbonyl O (red) and aromatic C-H (yellow)

A more in-depth analysis of hydrogen-bonded networks is provided by the Hydrogen Bond Propensity [155] tool, which, using the knowledge of hydrogen bonds in the CSD, predicts which donors and acceptors are likely to form hydrogen bonds in a crystal structure. The approach uses a logistic regression technique to provide the propensity (or probability) of a potential hydrogen-bonding network and has a potential application in identifying both likely and unusual hydrogen bonding. This can help to rationalise stable and metastable crystalline forms and assist in polymorph prediction.

3.3.4 Database Creation

The power of collecting data and aggregating it into databases and deriving knowledgebases has been clearly outlined above. The community is at a point where technological advances in crystallographic data generation mean there is potential for crystal structures to be generated at a much faster rate, which would in turn drive exploitation of databases. However, the process of growing a database involves collecting, cleaning, validating and adding value to data – the latter involving adding some chemical meaning to a crystal structure represented in CIF. Until recently this was a very time-consuming and manual process. DeCIFer [156] was therefore a necessary piece of software, developed to assign chemistry to crystallographic data, and has been used as part of the process of building the CSD since 2013. The process includes the detection of bonds, selection of a polymer unit where necessary, resolution of disorder and assignment of bond types and formal charges. Each assigned structure is accompanied by an estimate of the reliability that the chemistry has been correctly assigned and, in cases where the reliability is low, pointers towards the likely source(s) of errors. As well as information provided in an individual CIF, the collection of data in the CSD is used in various ways. Element-element pair distances from the CSD are used to aid in the detection and assignment of bonds. A Bayesian, probabilistic approach is used to improve the reliability and correctness of structure assignment and in some cases, existing chemical diagrams in the CSD are used in the generation of diagrams for incoming structures.

Adding value to data at the point of ingestion is only one aspect of this process – cleaning and validation are also key. In preparation for inclusion into the CSD, a piece of software, PreQuest [157], has been in use both at the CCDC and also by some crystallographers as the method of curating structures into a database. A recent successor to PreQuest, CSD Editor, enables Scientific Editors at the CCDC more easily to curate entries for the CSD and forms part of the database management infrastructure at the CCDC known as CSD-Xpedite. As a 'self-service' solution that enables crystallographers to better manage depositions, a new service 'My Structures' [158] was developed. My Structures partly addressed the data bottleneck and reduces burdens on both the part of the depositor and the database provider, and it enables users to view their previously deposited data, visualise structures, check the status of depositions and download and retrieve their datasets. My Structures and CSD Editor also provide the basis for users to create and maintain databases of their own structures. This means that in-house or proprietary crystal structures can be archived and searched either independently or in conjunction with the CSD.

3.3.5 Biological Integration

Small molecule crystal structures and organic ligands can be used to support research in areas of biology. Since the early 1990s, data derived from the CSD has been used

to aid protein structure refinement [159], proving useful in understanding interactions, in molecular geometry, in ligand fitting and in understanding overall protein structure [160]. Combining more precise geometric and derived data from small molecule structures in the CSD with PDB data can aid in the identification of potential new drug molecules. The following three areas illustrate how the overlap between small molecule and macromolecular crystallography can work together through: analysis of interactions in binding sites, searching of ligand-protein complexes and generating likely conformational binding poses for new potential drug molecules.

SuperStar [161] provides a knowledge-based approach to assist in identifying interaction sites in proteins. It uses information from IsoStar [146], which contains information about non-bonded interactions from both the CSD and PDB, to generate interaction maps within protein binding sites or around small molecules, i.e. it generates 'hot-spots' where a chosen interaction is particularly favourable.

Relibase [162] was a program building upon work published by Manfred Hendlich in 1998, which provides a means of storing and analysing protein-ligand complexes from the PDB. Query types included 2D substructure, 2D similarity, 3D substructure and sequence similarity searching. Relibase+ [163] additionally provided extensions to the software for handling crystallographic packing effects. Later publications provide further details of the design and curation of the Relibase database [164] and its application in drug design [165].

Another way that CSD data has been utilised is in the program GOLD [166] (Genetic Optimisation for Ligand Docking) which uses a genetic algorithm to predict the binding modes of flexible ligands into protein binding sites, a problem which is key in rational drug design. The docking of ligands is facilitated if they are first optimised using the CSD Conformer Generator [144]. A library of ring conformations extracted from the CSD can also be utilised by GOLD. CSD data has been used to parameterise other force fields, like COMPASS [167] and CHARMM [168], as well as in external conformer generators such as CONFECT [169] and BCL::CONF [170].

The value of being able to integrate different data sources has become clear in recent years, and to this end, simultaneous searching of the CSD and PDB has been integrated into CSD-CrossMiner [171]. This is a tool for pharmacophore-based searches which grew out of a collaboration between CCDC and the Computer-Aided Drug Design Section of F. Hoffmann-La Roche Ltd. at Roche Innovation Center Basel. It provides interactive searching for, amongst others, protein-ligand interaction patterns, ligand scaffolds or protein environments and can be used concurrently to search protein-ligand binding sites from the PDB and small organic molecules from the CSD using the same pharmacophore query. Searching just the CSD for 3D features from the binding site can reveal potential molecules that may bind to the protein. Reversing the search to look for proteins from the PDB may highlight other proteins with a similar binding pocket with which the ligand could also interact.

3.4 Areas Where Data-Driven Methods Are Making an Impact

The CSD can not only be used to search for structures and evaluate new structures but can also be used as a diverse experimental test set for computational algorithms, which seek to learn trends from the data and therefore predict molecules along with their useful properties. Machine learning approaches have received growing interest in recent years. Applications of the increasing volume of data in the CSD have been extensively covered in the recent review article by Taylor and Wood [15]. The single biggest research area where applications of the CSD are used is the pharmaceutical industry; however, there are many other areas where the database has been used to further research. The following section highlights a few examples of areas where crystallographic data from the CSD has been used to aid research.

3.4.1 Pharmaceuticals

The drug discovery journey begins with the search to find a molecule that has the correct shape and intermolecular interactions to bind in the target active site of a protein. The CSD can be used not only to search for suitable molecules but can also be exploited as a database of experimental conformations. These geometric preferences can be utilised when generating potential molecules in determining the most likely conformations [145, 169]. Intermolecular interaction data from the CSD is used in evaluating the interactions between the ligand and the binding site, especially to identify where changes to the molecule can be made to optimise the binding or physical properties of the molecule. Two approaches to drug candidate optimisation are bioisosteres, chemical substituents or groups with similar physical or chemical properties which produce broadly similar biological properties, and scaffold hopping, the search for compounds containing different core structures which retain similar activity due to the 3D location of binding groups [171–173].

In order to be able to develop an identified active molecule into a marketed drug, it is important to understand the solid form that will be given to the patient. This requires investigating a range of solid form characteristics, such as if the molecule is polymorphic or could have solvate or hydrate forms. While many techniques are utilised, the crystal structure plays an invaluable role in assessing a compound. The CSD has had a key role in underpinning collaborations with industry in the development of tools and methods to address these questions, firstly with the Pfizer Institute for Pharmaceutical Materials Sciences and then from 2008 the Crystal Form Consortium. A paper by Galek et al., celebrating the half millionth structure added to the CSD, is an example of the type of informatics that could be employed [174]. An in-depth analysis of the intermolecular interactions using Full Interaction Maps and Hydrogen Bond Propensity enabled an assessment of whether there are any unusual hydrogen bonds or unsatisfied hydrogen bond donor and acceptors in the crystal form. The presence of these may indicate the possibility of a more stable

form, which could warrant investigating experimentally. The assessment of the stability of a drug's solid form can be more complex if it is formulated as a co-crystal, an approach now often undertaken to improve a physical property of the compound. Using the data from the CSD on robust supramolecular synthons, a crystal engineering approach has been used to screen libraries of approved co-formers for co-crystal design. Understanding whether a drug molecule is likely to crystallise with a solvate or hydrate is still a challenge, and recent work has used machine learning in an attempt to predict the likelihood for formation [175, 176]. Machine learning approaches, using artificial neural networks, have also been the subject of work to ascertain other solid-state properties, e.g. melting point, lattice energy and crystal density [177].

Information from crystal structures is also valuable in assessing properties impacting pharmaceutical manufacturing, such as predicting the ease of compressing a potential drug form into a stable tablet. This method identifies slip planes between packed layers of molecules and considers the ease with which they could move over each other [178, 179]. The morphology of crystals has been shown to have an impact on the flowability of the material during manufacturing and is influenced by the growth rate of crystal faces [180]. This growth of a crystal face is influenced by the interactions between an attaching molecule and the molecules in the crystal, which has led to a variety of approaches to improve performance, by modifying the crystal habit, using different solvent systems, dopants and crystallisation conditions [181, 182].

3.4.2 Examples of Data-Driven Materials Discovery

MOFs have seen a dramatic rise this century [98], as shown above in Fig. 7. This interest in MOFs is partially due to their range of potential applications in gas storage, separation, catalysis, chemical sensors and drug delivery. MOF researchers have made extensive use of the CSD, and several groups have generated databases of structures, according to their own criteria, in order to then apply computational screening methods [183–185]. With these approaches in mind, a CSD subset of MOF structures, available through Conquest, has been assembled and is regularly updated as the CSD grows [97]. This subset can be used to characterise useful features such as classifying framework types [186], pore limiting diameter (PLD) and largest cavity diameter (LCD) [187], absorption, flexibility and other physical properties [188]. It has also been used to study the limits of hydrogen storage [189], gas adsorption [190] and separation of different gases [191, 192]. The key feature of all MOF materials is their high porosity, and there are now materials based on covalent organic frameworks (COFs), hydrogen-bonded frameworks (HOFs) and molecular cages that exhibit this high porosity. The exchange and identity of guest molecules in the pores of these materials has been of interest, and one application involving the encapsulation of guest molecules in the framework or cage has allowed their structure determination when it has not been possible to crystallise them independently, known as the crystalline sponge method [151, 193].

A number of research groups have data-mined the CSD and ICSD for relevant structures in their area of interest [194], including ferroelectrics [195] and non-linear optical materials [196]. As an example, one study extracted unbound 2,2′-bipyridine ligands and used atom distances to predict if the complex would exhibit spin crossover behaviour [197]. Another study performed property descriptor calculations on structures taken from the CSD and used this data to identify potential new organic semiconductors [198, 199]. A similar methodology was also used to find new a class of molecules for dye-sensitised solar cells [200].

It is not immediately and intuitively obvious that solid-state crystal structures should be relevant in developing soft matter such as gels. However, two different approaches have been used in this field. Firstly, a crystal engineering approach mined the CSD for hydrogen-bonded synthons for the design of new low molecular weight gelators [201]. Secondly, crystal morphology prediction has been used as an indication of the directionality of strong intermolecular interactions to develop a gel sensor to detect lead in paint [202].

Co-crystal design, as mentioned in an earlier section, is not only used in pharmaceutical and agrochemical research but can also be used in the rational design of energetic materials. It has additionally been shown, as a proof of concept, that it is possible to predict some key properties for energetic materials by machine learning [203]. All the structures in the CSD, and indeed crystallography itself, rely on being able to obtain a crystalline sample of the compound. This can be challenging for some areas of chemistry. Using compounds known to crystallise, due to a crystal structure being in the CSD, machine learning algorithms have been applied to determine if a molecule will crystallise at all [204].

4 Closing the Loop and Future Prospects

4.1 How Is Data Now Driving the Scientific Process and What Is the Future?

Many of the ‘solid-state rules’ have been worked out from amassing large volumes of data. This has resulted in the fact that molecular geometry prediction and assessment can readily be performed with a high degree of confidence, e.g. Mogul. In turn, in recent years there has been an explosion of crystal engineering resulting in a depth of understanding of intermolecular interactions that readily allows for design and control of relatively large solid-state structures – and this has led to fields such as supramolecular chemistry where complicated architectures and molecular interactions can be controlled so that these systems can now exhibit quite advanced functionality.

4.1.1 Data Science

There is still new ground to be covered and many grand challenges exist. For example, it is not readily possible for engineering and manufacturing sectors to get an immediate response to a demand for a material with a particular property or function. The future of structural science will be in addressing these issues and data is the key – not specifically collecting more data, although that is part of the solution in some respects, but more being able to link structural data to property/function data. Also, more complex hierarchical materials and multiple component or hybrid materials require a greater understanding to generate, so-called Directed Assembly – which is the subject of a 'Grand Challenge' supported by the UK Engineering and Physical Sciences Council [205]. In fact, there is an equal challenge in Directed Disassembly, which requires a deep level of understanding of structure, while we still fundamentally understand very little about the processes of nucleation and crystallisation. There is much data engineering to be done before the promised riches of data science can be realised – however, once it is possible to extract data from different databases on demand and run algorithms over them, then whole new research opportunities are opened up. Discovery, recognition and utilisation of patterns in data are fundamental in data science, and chemical/solid-state structure would be a key element in driving these approaches in application to chemical problems.

Structural similarity and structural informatics in the solid state are still relatively unexplored, yet have huge potential. This way of thinking raises the question of how big 'achievable solid-state space' actually is, i.e. if there were no barriers, how many crystal structures could we actually collect? This in turn leads on to recognition that there are currently significant 'gaps' in our databases. These gaps are mainly due to the fact that the majority of data arises from the traditional literature and therefore, e.g. collection of homologous series, that which is not deemed worthy of publication, 'uninteresting chemistry', etc. is work that is not undertaken. Data gathering exercises need to be given more value and credit if they are to be incorporated into a collection that can then be further utilised in many, many different ways. So, an immediate question is therefore one of how to identify these gaps and which ones are the most valuable to fill?

4.1.2 Higher-Resolution Structural Information

A logical progression that arises from the advances in instrumentation that are being realised is one of the resolutions of structural information that is potentially achievable. It is conceivable that with the right developments, then the time and effort taken to collect data and refine multipole models for charge density-level resolution will be drastically reduced. So, what is to stop this from becoming the normal approach to service crystallography? Databases that go beyond utilising atomic coordinates and allow investigation of electron distribution would be very powerful and open up new

levels of research. Furthermore, this level of resolution enables accurate quantum mechanical calculations based on experimental observations. So-called quantum crystallography [206] can calculate a range of properties and in particular quantitative energy calculations on interactions in the solid state. Such energetic calculations provide a new dimension for crystallography, as traditionally intermolecular interactions can only be inferred between atomic centres and quantified by a distance, but with this approach they can be directly observed. Furthermore, it is also possible to derive and deconvolute whole molecule-molecule interaction energies which begins to provide some of the information required to understand assembly (and disassembly) in the solid state.

Current state-of-the-art instrumentation is essentially at the point where this goal would be achievable. However, there are two further factors which would need to be addressed to achieve it. Firstly, there would have to be a significant input into developing accessible and sustainable software to process and refine this higher-resolution data and multipole models. Perhaps more of a challenge would be the need to effect a cultural change towards conducting such higher-level experiments, along with the significant amount of retraining and education that would be necessary. Furthermore, if there were to be a transition to collecting higher-resolution structures, then a question would be raised as to how to treat the one million structures already amassed in order to have comparable data.

4.1.3 Crystal Structure Prediction (CSP)

It is not the purpose of this review to comprehensively cover CSP; however, we note here the impact of inclusion of experimental data. In 2000 the CCDC ran its first blind test [207] to evaluate the state of methods for predicting crystal structures. The ability to predict more complex structures with greater accuracy has improved greatly in the subsequent blind test exercises. However, in the most recent, sixth blind test [208], CCDC entered itself for the first time and used the known structures in the CSD to predict unknown structures. The method used shape and packing similarity to generate potential crystallographic lattices. Although not a complete CSP solution, it proved valuable as a complementary technique reducing the chemical space needed to be searched using relatively cheap computer power and aid in structure ranking – CSP Speculator [209].

Recent developments from the CSP community are using machine learning approaches to explore the energy landscape of ensembles for predicted crystal structures [210]. While these approaches are currently more confined to understanding conformations of lowest energy, it would be entirely possible to combine experimental data and thereby use the power of both approaches. While CSP contributes significantly to the corpus of knowledge on crystal structures, it is also noteworthy in that it calculates a range of solid-state properties while doing so (or can be combined with computational property calculation), and therefore this complementary approach is destined to become increasingly powerful in the future.

4.1.4 Controlling Solid Form

Solid form control is the ability to select an appropriate morphology, structural polymorph and physical properties for a particular formulation of a particular compound and is one of the major challenges in pharmaceutical development. Many approaches, including informatics tools, are used in an attempt to reduce the risk of issues being found further along the development pipeline.

An example of one current direction is the use of multi-component systems (co-crystals, solvates and hydrates) which are an increasingly important consideration in the formulation of active pharmaceutical ingredients [211] and are now becoming available on the market. The void calculation functionality and visualisation of solvate and hydrate pockets or channels in Mercury led to the development of the Hydrate and Solvate Analysers to understand the behaviour of these molecules in crystal lattices. A study of molecular complementarity [212] used the CSD to describe pairs of molecules that form co-crystals with each other in terms of their calculated molecular properties. This has applications in the rational design of co-crystals and the modification of solid form's physical properties.

4.1.5 Crystallographic Data Driving Other Forms of Structure Solution and Refinement

The experimental approach and large volumes of rich data that we have discussed herein are founded in single-crystal X-ray diffraction analysis. While structure solution from this technique has some problematic examples, it is largely well understood, follows a well-trodden process, is universally accepted and has a 'data currency'. These factors are far from established in other techniques that probe the solid state. For example, solid-state structure determination from powder diffraction, pair distribution function, NMR crystallography and electron diffraction data are inherently challenging. All of these techniques draw from and leverage data and knowledge derived from acquisition of many single-crystal structures – from providing a starting point for refinement to imposing geometrical restraints/constraints derived from many observations.

Ab initio structure solution from powder diffraction data has been relatively widely possible for about 15 years and provides a good example of how acquired data can be used to influence and inform the process. The other solid-state structure determination approaches also use crystal structure data in much the same way – to facilitate matching calculated and observed data. There are a number of structure solution packages for analysing molecular structures from X-ray powder diffraction data, e.g. EXPO [213], GEST [214], GSAS-II [215] and TOPAS [216], including one provided by the CCDC [217], DASH [218, 219], which we use as an example here. DASH is a graphical user interface-driven program for solving crystal structures from measured powder diffraction data. It uses a simulated annealing approach to search for the global minimum in the agreement between observed and calculated

structure factors. Crystal structure data, from the CSD, can be used to steer this process. Its primary use is to reduce the search space, by the use of torsion angles restraints using data from Mogul for the conformational preference or by providing a starting geometry from a similar known single-crystal structure.

Probably the most important use of the CSD when using these techniques where 'direct' observation of the result is not possible is in the validation of the final structural parameters. DASH provides an exemplary use of this approach [220–224].

4.2 How Far Can Single-Crystal Diffraction Structure Analysis Be Developed?

4.2.1 Instrumentation

We consider first the future possibilities regarding synchrotron-based service crystallography and then go on to discuss what bearing these developments would have on home laboratory facilities.

Essentially there is nothing stopping synchrotron-based service crystallography being taken to the limit, which is the fully autonomous beamline. Such a facility would operate in a service provision mode, acting in a similar philosophy to many departmental institutional facilities. A user would send pre-mounted crystal and receive back datasets, without having any interaction with the beamline. Macromolecular crystallography already uses optical recognition to find the mounting loop and automatically centre it. If the crystal is essentially in the correct place, then it would be automatically centred, or it would be possible to define a region in which the crystal could be and search that region. A search could be performed optically, or by diffraction, or starting with a large beam, or by fluorescence/absorption for samples containing heavier elements. Once located, a second check could be run by collecting some data and determining the unit cell. This would enable determination of the x,y,z positional errors, which can then be fed back through the motorised positioner on the goniometer, further improving the centring. The next question to be posed would then be 'what are the data collection criteria for a crystal and how would we code them for a machine to use?' The first point to address would be whether there are any diffraction peaks at all. If there are, then the following would need to be assessed: Is it a diffraction pattern from a single crystal? What resolution does it extend to? Are the peak shapes good? During the data collections, are there, or will there be, signs of radiation damage?

Traditionally one would optically select the best crystal from a sample under the microscope and collect data on solely that crystal – if it produces good enough data, then the study is complete. However, there is often the situation where a screened crystal produces an acceptable diffraction pattern but that would ideally be better, in which case one has the dilemmas of: Are there any better crystals? Should the crystal be switched and is there anything of worth left in the sample? If the crystal is switched and subsequent attempts turn out worse, will the original crystal still be in a

good enough condition to be put back on the instrument? This leads to the conclusion that it would be better to collect data on several crystals and select the best result. The main factor against this approach is time: collection of data on multiple crystals takes time; processing and working up the data takes time. But if a dataset only takes a couple of minutes to acquire, why not collect on multiple crystals? Again, the factor against this is the time taken processing and evaluating all the datasets to find the best one, but if there was an algorithm that could process the data and then flag the best one to work on, then the problem is solved. If multiple datasets are collected, there is also the possibility of merging and scaling them into a single dataset with better averaging statistics and the ability for one collection to cover the shortfall of another. Data collection on multiple crystals also facilitates assessment for radiation damage. Some initial crystals could be sacrificed to determine the experimental conditions that would enable good data to be successfully collected. Approaches to mitigating radiation damage are likely to be specific and dependent on the mechanism of decay, but in some circumstances, methods of reducing the dose could include a better balance of decrease in beam intensity versus longer counting times, counting for the same time over multiple sweeps of the same strategy and summing the frames together, merging datasets from multiple crystals, reducing the beam size to be much smaller than the crystal and translating the crystal during the data collection so the same area is not continuously illuminated.

An increased speed of data collection also means multiple spheres of data may be collected. This enables more optimal performance of corrections programs such as SADABS and for the data to be finer sliced, e.g. 0.1° frame widths rather than 0.3°. In-house experiments (ALS) indicate that for a combination of SAINT [225] and SADABS on a compound with a small unit cell (all axes <13 Å), the data correction performs much better when narrower slices are used. There is also an improved signal-to-noise ratio when using narrower frame widths. If one considers the case of a 0.2° wide diffraction peak, for a 1° frame width and 1 s exposure, the background accumulates for 1 s while for the peak itself only 0.2 s; if collected with a 0.2° frame width for 0.2 s, the background accumulates for the same time as the peak.

With the move away from commercial instruments to custom ones, there is an opportunity to better and fully integrate them with the beamline. This allows for the operation of feedback loops such as the following: if the detector saturates, attenuation can be added or counting time decreased; if the dynamic range of the detector is being optimally and maximally under attenuation, this can be removed or the counting time increased. Additional data could also be collected, such as a fluorescence spectrum to provide qualitative element analysis for metals. This would also allow automated resonant scattering experiments to be run, making them as easy as a routine data collection. All the metadata for the collection can also be collated with beamline data such as wavelength, beam positions and intensity after each of the optical elements on the beamline – making troubleshooting more straightforward.

In the future the ideal detector would be an energy-resolving large area pixel detector. For monochromatic data collection, this would allow fluorescence background to be discriminated from diffraction and removed, improving the signal-to-noise ratio which is crucial for weakly diffracting crystals. More importantly it

would enable Laue diffraction techniques to be readily used for small molecule structure determinations on crystals with unknown unit cells. A single Laue diffraction frame contains more structural data than that of a whole monochromatic dataset, but with this technique, indexing an unknown unit cell is extremely difficult as while the position of the diffraction peak is known, the wavelength is not. With such a detector, it would be possible to determine both the position and wavelength, and thereby indexing the unit cell becomes straightforward. Laue diffraction would be ideal for very small crystals in small beams (where rotation could mean the crystal moving out of the beam) or in a sample environment with restricted access.

The implications of increased automation and advanced instrumentation are a little more subtle for the home laboratory. The most significant advance in a generation is the introduction of Hybrid Pixel Detectors, and these are going to become much more commonplace in the next decade. Time saving is significant, which for service crystallography means that turnaround times become negligible and a facility can operate more efficiently. The other main observation is that the data quality is significantly improved due to a better signal-to-noise ratio and lower background levels. It is questionable whether automation will have such a significant impact in the home laboratory, but some elements of this practice, such as rigorous collection of metadata, would mean that a better-quality result is reported. However, if the principles of automation were applied to the screening process, it would allow for the best possible result to be gleaned from a sample, as many crystals can be assessed and the best selected. This would ensure that the best possible result would come out of the home laboratory, and it would be possible to extend the current capability of the technique through approaches such as merging of multiple datasets, but the advances in the home laboratory in this respect will probably be rather more incremental.

The main question is one of whether full automation is really the best direction to take. On the one hand, it is not necessary to train and retrain users, saving time; users can no longer make mistakes during data collection, so the data quantity will improve; time will be freed up for difficult and time-consuming experiments. On the other hand, a fully automated beamline and even lab diffractometer system mean it is easy for anyone to collect data as long as they can mount the crystal in the correct position on the mounting loop. But we must be aware knowledge is being lost by this approach as these systems become black boxes and the understanding of the process is lost. Add to this automated refinement and over generation of students and the art of chemical crystallography will be lost as the fundamental knowledge of the experiment resides with fewer and fewer individuals.

4.2.2 Computing

Alongside instrumentation developments, in the modern era, it is now imperative to simultaneously consider a fully integrated and end-to-end software infrastructure to complement and support experimentation. It is important to recognise that such an infrastructure now needs to cover three distinct aspects: data acquisition and

processing, structure solution and refinement and integration with follow-on (data science) methods. The automation of data acquisition and structure solution discussed herein is a clear step in this direction that has largely been taken already. However, these approaches do not generally work for more complicated and difficult cases.

Better software and algorithms for data processing will be crucial for the future of the discipline. We have highlighted above that there is some work to be done as modern detectors evolve. However, current focus is on static structures and hence getting the best data integration for Bragg peaks – and in a significant number of cases, this is a simplistic view of the actual crystalline state. The ability to understand the total scatting pattern reveals the full behaviour in the crystal and provides great insights. This is not currently an easy or routine approach to take, and addressing this issue would open up a myriad of new structural chemistry. Routine analysis of the total scattering pattern would enable true modelling of 'disorder' and better analysis of local structure in complex materials and would provide the basis for dynamic crystallography methods to really thrive.

These advances in data processing call for advances in structure refinement. It will be necessary to develop new approaches to disorder modelling, and these could also be augmented by a closer operation with the databases. If refinement software worked more in tandem with the databases, it would be possible to learn from, and use, models (or parts of models) that already exist. A full and automated integration between the two would ensure that as the databases grow, they could increase in quality. Machine learning methods are now beginning to gain a lot of traction across many data-driven research areas – they have the potential to make structure refinements better, to make database records better and to power entirely new research in structural chemistry and in linking to other areas of science. There are also clear advantages in the convergence of experimental crystallography with Crystal Structure Prediction – the seamless interplay of these two approaches would mean that many more insights into an experimental structure would readily be possible and could feed into experiment design as well as interpretation.

4.2.3 Data

A Data Infrastructure

For a fully integrated and end-to-end software infrastructure to be realised, there is the necessity for the parallel development of a better, complementary data infrastructure. The crystallographic community has pioneered in many respects through the development of the Crystallographic Information Framework, and this provides the basis for a twenty-first-century data infrastructure. Modern automation approaches are a good example of leveraging data and metadata standards alongside software development. However, there are several areas where the data infrastructure clearly needs to be extended.

The main message of this review has the implication that there should be much more integration between the experiment and the database. With the correct infrastructure, a far greater number of structures could be included in the databases. For example, a lot of data mining and machine learning studies merely need to know that a structure exists, and its quality is of secondary importance – with a relatively small development and a change of publishing mindset, a vast number of new types of structures could be incorporated into databases. This 'data-fit-for-purpose' concept has the potential to hugely transform the amount of data available for follow-on studies; however, it requires further development of validation procedures and more accurate classification in databases, so that the integrity of a good-quality (subset) collection is not compromised for other areas of study. These developments could enable a clear, simple, fast, automated route from diffractometer to database which could in turn empower data mining, statistical and machine learning methods. These approaches are powerful not only in terms of performing structural chemistry analyses but also in making connections and correlations with data from other disciplines. There is a different data infrastructure that is required for this kind of work, and it is necessary to develop this for our subject. Currently data science involves a significant amount of data cleaning and transformation before the techniques can be applied – and this is often a very manual process. In order to make data science involving crystal structures seamless, it will be necessary to understand and implement new ways to interface between data collections – from both a metadata and descriptor perspective as well as via scripting and automation, e.g. via Application Programming Interfaces (API). Furthermore, changing the nature of the interactions between laboratory and database and data re-user will drive other developments, for example, the use of the database more interactively for structure refinement. It will also enhance integration between collaborators, complementary techniques and facilities.

Finally, it is necessary to extend the data infrastructure to enable greater inclusion of raw data. Currently there is no culture of sharing raw data in chemical crystallography – and in fact there is even a significant diversity in which individual facilities manage their own raw data. This lack of comprehensive management leads to difficulties in accessing the data (locally or globally) over time. Many other disciplines now routinely make raw data available, and there is an increasing pressure from funders and other stakeholders for this to be routine practice. There are many cultural, political and financial barriers to overcome for this to happen, but also some technical matters around description, validation, quality and storage would have to be addressed. Furthermore, it is worth considering whether it is necessary to make ALL raw data available, e.g. would it be necessary in the case where diffraction was very clean and all Bragg data had been accounted for? The IUCr CommDat [60] are considering these matters, and as a first step, the development of a CheckCIF-type utility for raw data is being investigated. Nevertheless, it would be very beneficial for a data infrastructure to support raw data where:

- Others, or future developments, could do a better job – in theory models could automatically be updated if better processing software were available.

- Greater scientific insight could be gained – e.g. secondary study of diffuse scattering purposefully not accounted for in the original model.
- It is necessary to support the review of claims made in publication – this would in turn support the data-fit-for-purpose concept (see above).
- Software developers could have a range of examples to develop their algorithms and code.
- Valuable examples arise that can be used for training in advanced data processing techniques.

With a more developed attitude to post processing, new approaches will also emerge. One development that would be particularly useful to improve quality and support dynamic crystallography is the practice of merging data collections to get a better composite dataset, e.g. in an approach to that used in macromolecular and serial crystallography communities.

The Data Landscape

'Chemical space' is a concept used by data scientists and cheminformaticians that refers to the property space spanned by all possible molecules, or compounds, for a given property. Chemical space is very large, e.g. pharmacological chemical space is estimated to cover 10^{63} molecules [226], and this even has many restrictions, e.g. does not include molecular weight >500 and only includes simple atoms (C, H, O, N, S) – many of the compounds are yet to exist. The Chemical Abstracts Service [227], which extracts compounds from the scientific literature, contains 158 million entries.

In comparison there are somewhat less than two million structures in the space covered by crystallographic databases. There are a number of reasons why these levels are so different – primarily that only a subset of materials are crystalline. However, a huge contributing factor is that most small molecule structures are determined as a service for synthesis chemists and subject to academic publishing rules in order to be available. There are further factors related to this phenomenon, e.g. that parameters/variables for crystallisation are not comprehensively tested; not all structures of a homologous/related family of compounds are deemed necessary (only one representative compound is suitable for publication – others may/may not have been determined); a structure may not be of suitable quality for publication; some compounds are 'not academically interesting enough' (from the perspective of a funder or synthesis chemist).

For these, and other, reasons, the crystallographic databases could be considered to have 'gaps' in them – particularly from the perspective of a data scientist. For some data science research to be valid or achievable, these gaps would have to be filled. There is a need for a change of mindset so that some of these gaps can begin to be filled – if the incentives were different, there is no reason why these gaps could not be identified and this used to drive synthesis programs and change culture. If a compound has been synthesised and there is ready access to crystallography, then

why should not its crystal structure be determined in order to contribute to the body of knowledge? Furthermore, it should be possible to make much more use of Crystal Structure Prediction as a complementary technique in order to fill the gaps. Such initiatives would enable more rigorous data mining and the use of machine learning and artificial intelligence methods to interrogate and rationalise crystal structure space.

4.3 Conclusions and Challenges

We have drawn numerous conclusions and insights throughout the various sections of this review. In doing so, the aim has been to not only provide insight into how specific aspects of crystallography are currently developing but also show how they can increasingly interact or integrate with other areas. This increased interoperation will provide a much richer methodology and enable crystallography to be a key component in a broad range of research long into the future.

Many research areas consider current times to be the Fourth Paradigm of Discovery [228] – that is one of data-intensive scientific discovery (or data science). The main message of this review is that chemical crystallography is very much operating in this regime, but that it has the potential to do much more and by taking a more data-integrated, or even data-centric, approach, it can be a leader in chemical and materials science research. The main challenge we face in achieving this is that it will involve changing working practice. The discipline needs to embrace different mindsets for working in the laboratory, new approaches to 'publishing' and new data science ways of undertaking research. Effecting this change will not only require new mindsets and approaches but also for researchers to be trained with new skills.

References

1. Helliwell JR (1992) Macromolecular crystallography with synchrotron radiation. https://doi.org/10.1017/CBO9780511524264
2. Clegg W (2000) Synchrotron chemical crystallography. J Chem Soc Dalt Trans:3223–3232
3. Katrusiak A (2008) High-pressure crystallography. Acta Crystallogr Sect A Found Crystallogr 64:135–148
4. Tidey JP, Wong HLS, Schröder M, Blake AJ (2014) Structural chemistry of metal coordination complexes at high pressure. Coord Chem Rev 277–278:187–207
5. Zhang J-P, Liao P-Q, Zhou H-L, Lin R-B, Chen X-M (2014) Single-crystal X-ray diffraction studies on structural transformations of porous coordination polymers. Chem Soc Rev 43:5789–5814
6. Hatcher LE, Raithby PR (2014) Dynamic single-crystal diffraction studies using synchrotron radiation. Coord Chem Rev 277–278:69–79
7. Barnett SA, Nowell H, Warren MR, Wilcox A, Allan DR (2016) Facilities for small-molecule crystallography at synchrotron sources. Protein Pept Lett 23:211–216

8. Nowell H, Barnett SA, Christensen KE, Teat SJ, Allan DR (2012) I19, the small-molecule single-crystal diffraction beamline at diamond light source. J Synchrotron Radiat 19:435–441
9. McCormick LJ, Giordano N, Teat SJ, Beavers CM (2017) Chemical crystallography at the advanced light source. Crystals 7:382
10. Hursthouse MB, Coles SJ (2014) The UK national crystallography service; its origins, methods and science. Crystallogr Rev 20:117–154
11. Coles SJ, Gale PA (2012) Changing and challenging times for service crystallography. Chem Sci 3:683–689
12. EPSRC (2020) EPSRC national research facilities programme. https://epsrc.ukri.org/research/facilities/access/nationalresearch/
13. Hackert ML, Van Meervelt L, Helliwell JR, McMahon B (2016) Open data in a big data world: a position paper for crystallography. https://www.iucr.org/iucr/open-data
14. Groom CR, Bruno IJ, Lightfoot MP, Ward SC (2016) The Cambridge Structural Database. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 72:171–179
15. Taylor R, Wood PA (2019) A million crystal structures: the whole is greater than the sum of its parts. Chem Rev 119:9427–9477
16. Phillips GN (2015) The future of dynamic structural science. Crystallogr Rev 21:310–310
17. Coles SJ, Hursthouse MB (2004) Focusing optics for molybdenum radiation: a bright laboratory source for small-molecule crystallography. J Appl Crystallogr 37:988–992
18. He BB (2009) Two-dimensional X-ray diffraction. https://doi.org/10.1002/9780470502648
19. Hemberg O, Otendal M, Hertz HM (2003) Liquid-metal-jet anode electron-impact x-ray source. Appl Phys Lett 83:1483–1485
20. Otendal M, Tuohimaa T, Vogt U, Hertz HM (2008) A 9keV electron-impact liquid-gallium-jet x-ray source. Rev Sci Instrum 79:016102
21. Gruner SM, Tate MW, Eikenberry EF (2002) Charge-coupled device area x-ray detectors. Rev Sci Instrum 73:2815–2842
22. Allé P, Wenger E, Dahaoui S, Schaniel D, Lecomte C (2016) Comparison of CCD, CMOS and hybrid pixel x-ray detectors: detection principle and data quality. Phys Scr 91:063001
23. Kraft P, Bergamaschi A, Broennimann C et al (2009) Performance of single-photon-counting PILATUS detector modules. J Synchrotron Radiat 16:368–375
24. Rigaku (2020) Rigaku HyPix. https://www2.rigaku.com/products/detectors/hypix6000
25. Bruker (2020) Bruker PHOTON. https://www.bruker.com/products/x-ray-diffraction-and-elemental-analysis/single-crystal-x-ray-diffraction/sc-xrd-components/sc-xrd-components/overview/sc-xrd-components/detectors/photon-iii.html
26. Thompson AC, Westbrook EM, Lavender WM, Nix JC (2014) A large area CMOS detector for shutterless collection of x-ray diffraction data. J Phys Conf Ser 493:012019
27. Elder FR, Gurewitsch AM, Langmuir RV, Pollock HC (1947) Radiation from electrons in a synchrotron. Phys Rev 71:829–830
28. Robinson AL (2001) X-ray data booklet. In: Hist. Synchrotron Radiat. Lawrence Berkeley National Laboratory, p Section 2.2
29. Kim KJ (2001) X-ray data booklet. In: Charact. Synchrotron Radiat. Lawrence Berkeley National Laboratory, p Section 2.1
30. Attwood D (1999) Soft X-rays and extreme ultraviolet radiation. https://doi.org/10.1017/CBO9781139164429
31. Duke P. Synchrotron radiation
32. Marks S, Zbasnik J, Byme W et al (2002) ALS superbend magnet performance. IEEE Trans Appl Supercond 12:149–152
33. Eriksson M, van der Veen JF, Quitmann C (2014) Diffraction-limited storage rings – a window to the science of tomorrow. J Synchrotron Radiat 21:837–842
34. Raimondi P (2016) ESRF-EBS: the extremely brilliant source project. Synchrotron Radiat News 29:8–15
35. Inoue I, Osaka T, Tamasaku K, Ohashi H, Yamazaki H, Goto S, Yabashi M (2018) An X-ray harmonic separator for next-generation synchrotron X-ray sources and X-ray free-electron lasers. J Synchrotron Radiat 25:346–353
36. Kirkpatrick P, Baez AV (1948) Formation of optical images by X-rays. J Opt Soc Am 38:766

37. Helliwell JR (1984) Synchrotron X-radiation protein crystallography: instrumentation, methods and applications. Rep Prog Phys 47:1403–1497
38. Ballabriga R, Alozy J, Blaj G et al (2013) The Medipix3RX: a high resolution, zero dead-time pixel detector readout chip allowing spectroscopic imaging. J Instrum 8:C02016
39. Hatsui T, Graafsma H (2015) X-ray imaging detectors for synchrotron and XFEL sources. IUCrJ 2:371–383
40. Graafsma H, Becker J, Gruner SM (2018) Integrating hybrid area detectors for storage ring and free-electron laser applications. In: Synchrotron light sources free. Lasers. Springer, Cham, pp 1–31
41. Leonarski F, Redford S, Mozzanica A et al (2018) Fast and accurate data collection for macromolecular crystallography using the JUNGFRAU detector. Nat Methods 15:799–804
42. Harding MM, Kariuki BM, Cernik R, Cressey G (1994) The structure of aurichalcite, (Cu,Zn) 5 (OH) 6 (CO 3) 2, determined from a microcrystal. Acta Crystallogr Sect B Struct Sci 50:673–676
43. Cernik RJ, Clegg W, Catlow CRA, Bushnell-Wye G, Flaherty JV, Greaves GN, Burrows I, Taylor DJ, Teat SJ, Hamichi M (1997) A new high-flux chemical and materials crystallography station at the SRS daresbury. 1. Design, construction and test results. J Synchrotron Radiat 4:279–286
44. SmarAct (2020) SMARGON. https://www.smaract.com/smargon
45. Cosier BJ, Glazer AM (1986) A nitrogen-gas-stream cryostat for general X-ray diffraction studies. J Appl Crystallogr 19:105–107
46. Agilent (2020) CrysAlis PRO. https://www.rigaku.com/en/products/smc/crysalis
47. Johnson NR, Waddell PG, Clegg W, Probert MR (2017) Remote access revolution: chemical crystallographers enter a new era at diamond light source beamline I19. Crystals 7:360
48. Delageniere S, Brenchereau P, Launer L et al (2011) ISPyB: an information management system for synchrotron macromolecular crystallography. Bioinformatics 27:3186–3192
49. Allan D, Nowell H, Barnett S et al (2017) A novel dual air-bearing fixed-χ diffractometer for small-molecule single-crystal X-ray diffraction on beamline I19 at diamond light source. Crystals 7:336
50. Diamond Light Source (2020) DLS data management policy
51. ESRF (2020) ESRF data management policy
52. APS (2020) APS data management policy
53. Christensen J, Horton PN, Bury CS, Dickerson JL, Taberman H, Garman EF, Coles SJ (2019) Radiation damage in small-molecule crystallography: fact not fiction. IUCrJ 6:703–713
54. Garman EF, Weik M (2017) Radiation damage in macromolecular crystallography. In: Methods Mol Biol. pp 467–489
55. Royal Society (1665) Epistle dedicatory. Philos Trans R Soc Lond. https://doi.org/10.1098/rstl.1665.0001
56. Hall SR (1991) The STAR file: a new format for electronic data transfer and archiving. J Chem Inf Model 31:326–333
57. Hall SR, Allen FH, Brown ID (1991) The crystallographic information file (CIF): a new standard archive file for crystallography. Acta Crystallogr Sect A Found Crystallogr 47:655–685
58. Brown ID, McMahon B (2002) CIF: the computer language of crystallography. Acta Crystallogr Sect B Struct Sci 58:317–324
59. Bernstein HJ, Bollinger JC, Brown ID, Gražulis S, Hester JR, McMahon B, Spadaccini N, Westbrook JD, Westrip SP (2016) Specification of the crystallographic information file format, version 2.0. J Appl Crystallogr 49:277–284
60. IUCr (2020) Committee on data (CommDat). https://www.iucr.org/iucr/governance/advisory-committees/committee-on-data
61. Spek AL (2003) Single-crystal structure validation with the program PLATON. J Appl Crystallogr 36:7–13
62. Spek AL (2009) Structure validation in chemical crystallography. Acta Crystallogr Sect D Biol Crystallogr 65:148–155
63. Spek AL (2018) What makes a crystal structure report valid? Inorg Chim Acta 470:232–237

64. Helliwell JR, McMahon B, Guss JM, Kroon-Batenburg LMJ (2017) The science is in the data. IUCrJ 4:714–722
65. Royal Society of Chemistry (2020) CrystEngComm. https://www.rsc.org/journals-books-databases/about-journals/crystengcomm/
66. American Chemical Society (2020) Crystal growth & design. https://pubs.acs.org/journal/cgdefu
67. Bruno I, Gražulis S, Helliwell JR, Kabekkodu SN, McMahon B, Westbrook J (2017) Crystallography and Databases. Data Sci J. https://doi.org/10.5334/dsj-2017-038
68. Gražulis S, Chateigner D, Downs RT et al (2009) Crystallography open database – an open-access collection of crystal structures. J Appl Crystallogr 42:726–729
69. Hellenbrandt M (2004) The inorganic crystal structure database (ICSD)—present and future. Crystallogr Rev 10:17–22
70. Allen FH, Kennard O, Watson DG, Brammer L, Orpen AG, Taylor R (1987) Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds. J Chem Soc Perkin Trans 2:S1
71. Orpen AG, Brammer L, Allen FH, Kennard O, Watson DG, Taylor R (1989) Supplement. Tables of bond lengths determined by X-ray and neutron diffraction. Part 2. Organometallic compounds and co-ordination complexes of the d- and f-block metals. J Chem Soc Dalt Trans: S1
72. Coles SJ, Frey JG, Hursthouse MB, Light ME, Meacham KE, Marvin DJ, Surridge M (2005) ECSES – examining crystal structures using `e-science': a demonstrator employing web and grid services to enhance user participation in crystallographic experiments. J Appl Crystallogr 38:819–826
73. Digital Science (2020) Figshare. https://figshare.com/
74. Bird C, Coles SJ, Frey JG (2015) The evolution of digital chemistry at Southampton. Mol Inform 34:585–597
75. CCDC (2020) CSD communications. https://www.ccdc.cam.ac.uk/Community/csd-communications/
76. Wilkinson MD, Dumontier M, Aalbersberg IJ et al (2016) The FAIR guiding principles for scientific data management and stewardship. Sci Data 3:160018
77. GO-FAIR (2020) FAIR principles. https://www.go-fair.org/fair-principles/
78. Mons B, Neylon C, Velterop J, Dumontier M, da Silva Santos LOB, Wilkinson MD (2017) Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European Open Science cloud. Inf Serv Use 37:49–56
79. Coles SJ, Frey JG, Willighagen EL, Chalk SJ (2019) Taking FAIR on the ChIN: the chemistry implementation network. Data Intell:131–138
80. NSF (2020) National Science Foundation. https://www.nsf.gov/bfa/dias/policy/dmp.jsp
81. UKRI (2020) UK Research & Innovation. https://epsrc.ukri.org/about/standards/researchdata/
82. ERC (2020) European Research Council. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf%0D%0Ahttps://www.ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_e
83. FAIR Databases (2020) FAIR crystallographic databases. https://www.rcsb.org/pages/about-us/mission
84. Dolomanov OV, Bourhis LJ, Gildea RJ, Howard JAK, Puschmann H, IUCr (2009) OLEX2: a complete structure solution, refinement and analysis program. J Appl Crystallogr 42:339–341
85. Betteridge PW, Carruthers JR, Cooper RI, Prout K, Watkin DJ (2003) CRYSTALS version 12: software for guided crystal structure analysis. J Appl Crystallogr 36:1487–1487
86. Berman H, Henrick K, Nakamura H (2003) Announcing the worldwide protein data bank. Nat Struct Mol Biol 10:980–980
87. Kennard O (1996) From private data to public knowledge. Portland Press, London
88. Marsh RE, Schomaker V (1979) Some incorrect space groups in inorganic chemistry, volume 16. Inorg Chem 18:2331–2336
89. Henling LM, Marsh RE (2014) Some more space-group corrections. Acta Crystallogr Sect C Struct Chem 70:834–836

90. Fronczek FR (2018) The inverse Marsh error. Acta Crystallogr Sect A Found Adv 74:a60–a60
91. Morinaka Y, Sato S, Wakamiya A, Nikawa H, Mizorogi N, Tanabe F, Murata M, Komatsu K, Furukawa K, Kato T, Nagase S, Akasaka T, Murata Y (2013) X-ray observation of a helium atom and placing a nitrogen atom inside He@ C 60 and He@ C 70. Nat Commun 4:1554. https://doi.org/10.1038/ncomms2574
92. Silver MA, Cary SK, Johnson JA, Baumbach RE, Arico AA, Luckey M, Urban M, Wang JC, Polinski MJ, Chemey A, Liu G, Chen K-W, Van Cleve SM, Marsh ML, Eaton TM, van de Burgt LJ, Gray AL, Hobart DE, Hanson K, Maron L, Gendron F, Autschbach J, Speldrich M, Kögerler P, Yang P, Braley J, Albrecht-Schmitt TE (2016) Characterization of berkelium (III) dipicolinate and borate compounds in solution and the solid state. Science 353:888. https://doi.org/10.1126/science.aaf3762
93. Apostolidis C, Schimmelpfennig B, Magnani N, Lindqvist-Reis P, Walter O, Sykora R, Morgenstern A, Colineau E, Caciuffo R, Klenze R, Haire RG, Rebizant J, Bruchertseifer F, Fanghänel T (2010) $[An(H_2O)_9](CF_3SO_3)_3$ (An=U-Cm, Cf): exploring their stability, structural chemistry, and magnetic behavior by experiment and theory. Angew Chem Int Ed 49:6343. https://doi.org/10.1002/anie.201001077
94. Polinski MJ, Garner III EB, Maurice R, Planas N, Stritzinger JT, Gannon Parker T, Cross JN, Green TD, Alekseev EV, Van Cleve SM, Depmeier W, Gagliardi L, Shatruk M, Knappenberger KL, Liu G, Skanthakumar S, Soderholm L, Dixon DA, Albrecht-Schmitt TE (2014) Unusual structure, bonding and properties in a Californium borate. Nat Chem 6:387. https://doi.org/10.1038/nchem.1896
95. Cary SK, Vasiliu M, Baumbach RE, Stritzinger JT, Green TD, Diefenbach K, Cross JN, Knappenberger KL, Liu G, Silver MA, DePrince AE, Polinski MJ, Van Cleve SM, House JH, Kikugawa N, Gallagher A, Arico AA, Dixon DA, Albrecht-Schmitt TE (2015) Emergence of californium as the second transitional element in the actinide series. Nat Commun 6:6827. https://doi.org/10.1038/ncomms7827
96. Cary SK, Su J, Galley SS, Albrecht-Schmitt TE, Batista ER, Ferrier MG, Kozimor SA, Mocko V, Scott BL, Van Alstine CE, White FD, Yang P (2018) A series of dithiocarbamates for americium, curium, and californium. Dalton Trans 47:14452. https://doi.org/10.1039/C8DT02658K
97. Moghadam PZ, Li A, Wiggin SB, Tao A, Maloney AGP, Wood PA, Ward SC, Fairen-Jimenez D (2017) Development of a Cambridge Structural Database subset: a collection of metal–organic frameworks for past, present, and future. Chem Mater 29:2618–2625
98. Furukawa H, Cordova KE, O'Keeffe M, Yaghi OM (2013) The chemistry and applications of metal-organic frameworks. Science (80-) 341:1230444
99. Tovee C, Ward S, Sarjeant A, Bruno I (2018) Reporting crystal structure data: recent insights. Abstr Pap Am Chem Soc 256
100. Van Der Sluis P, Spek AL (1990) BYPASS: an effective method for the refinement of crystal structures containing disordered solvent regions. Acta Crystallogr Sect A Found Crystallogr 46:194–201
101. Allen FH, Bellard S, Brice MD et al (1979) The Cambridge crystallographic data centre: computer-based search, retrieval, analysis and display of information. Acta Crystallogr Sect B Struct Crystallogr Cryst Chem 35:2331–2339
102. Johnson CK (1965) ORTEP: a fortran thermal-ellipsoid plot program for crystal structure illustrations
103. Crystal Impact, Putz H, Brandenburg K. Diamond – crystal and molecular structure visualization
104. CrystalMaker Software Ltd CrystalMaker®
105. Siemens (1994) XP
106. Watkin DJ, Prout CK, Pearce LJ (1996) Cameron
107. Johnson CK, Burnett MN (1996) ORTEPIII
108. Barbour LJ (2001) X-seed — a software tool for supramolecular crystallography. J Supramol Chem 1:189–191

109. Motherwell WDS, Shields GP, Allen FH (1999) Visualization and characterization of non-covalent networks in molecular crystals: automated assignment of graph-set descriptors for asymmetric molecules. Acta Crystallogr Sect B Struct Sci 55:1044–1056
110. Bruno IJ, Cole JC, Edgington PR, Kessler M, Macrae CF, McCabe P, Pearson J, Taylor R (2002) New software for searching the Cambridge Structural Database and visualizing crystal structures. Acta Crystallogr Sect B Struct Sci 58:389–397
111. Taylor R, Macrae CF (2001) Rules governing the crystal packing of mono- and dialcohols. Acta Crystallogr Sect B Struct Sci 57:815–827
112. CCDC (1994) Vista – a program for the analysis and display of data retrieved from the CSD
113. Sykes RA, McCabe P, Allen FH, Battle GM, Bruno IJ, Wood PA (2011) New software for statistical analysis of Cambridge Structural Database data. J Appl Crystallogr 44:882–886
114. Macrae CF, Edgington PR, McCabe P, Pidcock E, Shields GP, Taylor R, Towler M, van de Streek J, IUCr (2006) *Mercury*: visualization and analysis of crystal structures. J Appl Crystallogr 39:453–457
115. Allen FH, Davies JE, Galloy JJ, Johnson O, Kennard O, Macrae CF, Mitchell EM, Mitchell GF, Smith JM, Watson DG (1991) The development of versions 3 and 4 of the Cambridge Structural Database system. J Chem Inf Comput Sci 31:187–204
116. Stumpfe D, Bajorath J (2011) Similarity searching. Wiley Interdiscip Rev Comput Mol Sci 1:260–282
117. Cambridge Crystallographic Data Centre CellCheckCSD – The Cambridge Crystallographic Data Centre (CCDC)
118. White FJ, Gál Z, Griffin A, Skarzynski T, Meyer M, Prochniak G, Wood PA, Thomas IR (2011) A new interface to the Cambridge Structural Database (CSD) in CrysAlisPro. Acta Crystallogr Sect A Found Crystallogr 67:C404–C404
119. Chisholm JA, Motherwell S (2004) A new algorithm for performing three-dimensional searches of the Cambridge Structural Database. J Appl Crystallogr 37:331–334
120. Macrae CF, Bruno IJ, Chisholm JA et al (2008) *Mercury CSD 2.0* – new features for the visualization and investigation of crystal structures. J Appl Crystallogr 41:466–470
121. Gelbrich T, Hursthouse MB (2005) A versatile procedure for the identification, description and quantification of structural similarity in molecular crystals. CrystEngComm 7:324
122. Chisholm JA, Motherwell S (2005) COMPACK : a program for identifying crystal structure similarity using distances. J Appl Crystallogr 38:228–231
123. Rohlíček J, Skořepová E, Babor M, Čejka J, IUCr (2016) *CrystalCMP*: an easy-to-use tool for fast comparison of molecular packing. J Appl Crystallogr 49:2172–2183
124. Salbego PRSS, Bender CR, Hörner M, Zanatta N, Frizzo CP, Bonacorso HG, Martins MAPP (2018) Insights on the similarity of supramolecular structures in organic crystals using quantitative indexes. ACS Omega 3:2569–2578
125. Thomas IR, Bruno IJ, Cole JC, Macrae CF, Pidcock E, Wood PA (2010) WebCSD : the online portal to the Cambridge Structural Database. J Appl Crystallogr 43:362–366
126. FIZ Karlsruhe. https://icsd.products.fiz-karlsruhe.de/
127. RCSB PDB. https://www.rcsb.org/
128. Crystallography Open Database. http://www.crystallography.net/cod/index.php
129. CrystalWorks. https://cds.dl.ac.uk/cds/datasets/crys/cweb/cworks.html
130. PSDS (2020) Physical sciences data-science service. www.psds.ac.uk
131. Johnston D (2020) Symmetry resources at Otterbein University. https://symotter.org/
132. IUCr (2020) Educational web sites and resources of interest. https://www.iucr.org/education/resources
133. BCA (2020) Learn Crystallography. https://learn.crystallography.org.uk/
134. ECA (2020) Knowledge centre and resources. https://gig03eca.wixsite.com/gig03eca/knowledge-centre
135. CCDC (2020) Educational resources. https://www.ccdc.cam.ac.uk/Community/educationalresources/
136. EBI (2020) Teaching materials. https://www.ebi.ac.uk/pdbe/training/teaching-materials
137. PDB (2020) PDB-101. https://pdb101.rcsb.org/
138. CCDC (2020) CIF deposition and validation service. https://www.ccdc.cam.ac.uk/deposit

139. IUCr (2020) CheckCIF. http://checkcif.iucr.org/
140. Orpen AG, Brammer L, Allen FH, Watson DG, Taylor R (2006) Typical interatomic distances: organometallic compounds and coordination complexes of the d – and f -block metals. Int Tables Crystallogr C:812–896
141. Allen FH, Watson DG, Brammer L, Orpen AG, Taylor R (2006) Typical interactomic distances: organic compounds. Int Tables Crystallogr C:790–811
142. Bruno IJ, Cole JC, Kessler M et al (2004) Retrieval of crystallographically-derived molecular geometry information. J Chem Inf Comput Sci 44:2133–2144
143. Cottrell SJ, Olsson TSG, Taylor R, Cole JC, Liebeschuetz JW (2012) Validating and understanding ring conformations using small molecule crystallographic data. J Chem Inf Model 52:956–962
144. Cole JC, Korb O, McCabe P, Read MG, Taylor R (2018) Knowledge-based conformer generation using the Cambridge Structural Database. J Chem Inf Model 58:615–629
145. Taylor R, Cole J, Korb O, McCabe P (2014) Knowledge-based libraries for predicting the geometric preferences of druglike molecules. J Chem Inf Model 54:2500–2514
146. Bruno IJ, Cole JC, Lommerse JPM, Rowland RS, Taylor R, Verdonk ML (1997) IsoStar: a library of information about nonbonded interactions. J Comput Aided Mol Des 11:525–537
147. Taylor R (2016) It Isn't, it is: the C-H···X (X = O, N, F, Cl) interaction really is significant in crystal packing. Cryst Growth Des 16:4165–4168
148. Bauzá A, Seth SK, Frontera A (2019) Tetrel bonding interactions at work: impact on tin and lead coordination compounds. Coord Chem Rev 384:107–125
149. Bauzá A, Frontera A (2015) Aerogen bonding interaction: a new supramolecular force? Angew Chemie – Int Ed 54:7340–7343
150. Mikherdov AS, Kinzhalov MA, Novikov AS, Boyarskiy VP, Boyarskaya IA, Avdontceva MS, Kukushkin VY (2018) Ligation-enhanced π-hole···π interactions involving isocyanides: effect of π-hole···π noncovalent bonding on conformational stabilization of acyclic diaminocarbene ligands. Inorg Chem 57:6722–6733
151. Rissanen K (2017) Crystallography of encapsulated molecules. Chem Soc Rev 46:2638–2648
152. Wood PA, Olsson TSG, Cole JC, Cottrell SJ, Feeder N, Galek PTA, Groom CR, Pidcock E (2013) Evaluation of molecular crystal structures using full interaction maps. CrystEngComm 15:65–72
153. Feeder N, Pidcock E, Reilly AM, Sadiq G, Doherty CL, Back KR, Meenan P, Docherty R (2015) The integration of solid-form informatics into solid-form selection. J Pharm Pharmacol 67:857–868
154. Galek PTA, Pidcock E, Wood PA, Feeder N, Allen FH (2016) Navigating the solid form landscape with structural informatics. In: Computational pharmaceutical solid state chemistry. Wiley, Hoboken, pp 15–35
155. Galek PTA, Fábián L, Motherwell WDS, Allen FH, Feeder N (2007) Knowledge-based model of hydrogen-bonding propensity in organic crystals. Acta Crystallogr Sect B Struct Sci 63:768–782
156. Bruno IJ, Shields GP, Taylor R (2011) Deducing chemical structure from crystallographically determined atomic coordinates. Acta Crystallogr Sect B Struct Sci 67:333–349
157. CCDC (2020) PreQuest. https://www.ccdc.cam.ac.uk/solutions/csd-system/components/prequest/
158. CCDC (2020) My structures. https://www.ccdc.cam.ac.uk/support-and-resources/support/case/?caseid=a567fad5-20b7-e611-837e-00505686f06e
159. Cole JC, Giangreco I, Groom CR (2017) Using more than 801 296 small-molecule crystal structures to aid in protein structure refinement and analysis. Acta Crystallogr Sect D Struct Biol 73:234–239
160. Groom CR, Cole JC (2017) The use of small-molecule structures to complement protein-ligand crystal structures in drug discovery. Acta Crystallogr Sect D Struct Biol 73:240–245
161. Verdonk ML, Cole JC, Taylor R (1999) SuperStar: a knowledge-based approach for identifying interaction sites in proteins. J Mol Biol 289:1093–1108
162. Hendlich M (1998) Databases for protein–ligand complexes. Acta Crystallogr Sect D Biol Crystallogr 54:1178–1182

163. Bergner A, Gunther J, Hendlich M, Klebe G, Verdonk M (2001) Use of relibase for retrieving complex three-dimensional interaction patterns including crystallographic packing effects. Biopolymers 61:99–110
164. Hendlich M, Bergner A, Günther J, Klebe G (2003) Relibase: design and development of a database for comprehensive analysis of protein–ligand interactions. J Mol Biol 326:607–620
165. Günther J, Bergner A, Hendlich M, Klebe G (2003) Utilising structural knowledge in drug design strategies: applications using relibase. J Mol Biol 326:621–636
166. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727–748
167. Sun H, Jin Z, Yang C, Akkermans RLC, Robertson SH, Spenley NA, Miller S, Todd SM (2016) COMPASS II: extended coverage for polymer and drug-like molecule databases. J Mol Model 22:47
168. Vermaas JV, Petridis L, Ralph J, Crowley MF, Beckham GT (2019) Systematic parameterization of lignin for the CHARMM force field. Green Chem 21:109–122
169. Schärfer C, Schulz-Gasch T, Hert J, Heinzerling L, Schulz B, Inhester T, Stahl M, Rarey M (2013) Inside cover: CONFECT: conformations from an expert collection of torsion patterns (ChemMedChem 10/2013). ChemMedChem 8:1574–1574
170. Kothiwale S, Mendenhall JL, Meiler J (2015) BCL::Conf: small molecule conformational sampling using a knowledge based rotamer library. J Cheminform 7:47
171. Korb O, Kuhn B, Hert J, Taylor N, Cole J, Groom C, Stahl M (2016) Interactive and versatile navigation of structural databases. J Med Chem 59:4257–4266
172. Groom CR, Olsson TSG, Liebeschuetz JW, Bardwell DA, Bruno IJ, Allen FH (2012) Mining the Cambridge Structural Database for bioisosteres. In: Bioisosteres medicinal chemistry. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, pp 75–101
173. Cresset (2020) Fragments and conformations from the CCDC's Cambridge Structural Database accessible through Cresset's Spark. https://www.cresset-group.com/about/news/fragments-and-conformations-from-the-ccdcs-cambrid/. Accessed 4 July 2019
174. Galek PTA, Pidcock E, Wood PA, Bruno IJ, Groom CR (2012) One in half a million: a solid form informatics study of a pharmaceutical crystal structure. CrystEngComm 14:2391–2403
175. Takieddin K, Khimyak YZ, Fábián L (2016) Prediction of hydrate and solvate formation using statistical models. Cryst Growth Des 16:70–81
176. Xin D, Gonnella NC, He X, Horspool K (2019) Solvate prediction for pharmaceutical organic molecules with machine learning. Cryst Growth Des 19:1903–1911
177. Rama Krishna G, Ukrainczyk M, Zeglinski J, Rasmuson ÅC (2018) Prediction of solid state properties of cocrystals using artificial neural network modeling. Cryst Growth Des 18:133–144
178. Bryant MJ, Maloney AGP, Sykes RA (2018) Predicting mechanical properties of crystalline materials through topological analysis. CrystEngComm 20:2698–2704
179. Wang C, Sun CC (2019) Computational techniques for predicting mechanical properties of organic crystals: a systematic evaluation. Mol Pharm 16:1732–1741
180. Pudasaini N, Upadhyay PP, Parker CR, Hagen SU, Bond AD, Rantanen J (2017) Downstream processability of crystal habit-modified active pharmaceutical ingredient. Org Process Res Dev 21:571–577
181. Turner TD, Hatcher LE, Wilson CC, Roberts KJ (2019) Habit modification of the active pharmaceutical ingredient lovastatin through a predictive solvent selection approach. J Pharm Sci 108:1779–1787
182. Hooper D, Clarke FC, Docherty R, Mitchell J, Snowden MJ (2017) Effects of crystal habit on the sticking propensity of ibuprofen—a case study. Int J Pharm 531:266–275
183. Chung YG, Camp J, Haranczyk M, Sikora BJ, Bury W, Krungleviciute V, Yildirim T, Farha OK, Sholl DS, Snurr RQ (2014) Computation-ready, experimental metal–organic frameworks: a tool to enable high-throughput screening of nanoporous crystals. Chem Mater 26:6185–6192
184. First EL, Floudas CA (2013) MOFomics: computational pore characterization of metal-organic frameworks. Microporous Mesoporous Mater 165:32–39

185. Watanabe T, Sholl DS (2012) Accelerating applications of metal–organic frameworks for gas adsorption and separation by computational screening of materials. Langmuir 28:14114–14128
186. Barthel S, Alexandrov EV, Proserpio DM, Smit B (2018) Distinguishing metal–organic frameworks. Cryst Growth Des 18:1738–1747
187. Miklitz M, Jelfs KE (2018) pywindow: automated structural analysis of molecular pores. J Chem Inf Model 58:2387–2391
188. Coudert F-X, Fuchs AH (2016) Computational characterization and prediction of metal–organic framework properties. Coord Chem Rev 307:211–236
189. Goldsmith J, Wong-Foy AG, Cafarella MJ, Siegel DJ (2013) Theoretical limits of hydrogen storage in metal–organic frameworks: opportunities and trade-offs. Chem Mater 25:3373–3382
190. Moghadam PZ, Islamoglu T, Goswami S, Exley J, Fantham M, Kaminski CF, Snurr RQ, Farha OK, Fairen-Jimenez D (2018) Computer-aided discovery of a metal–organic framework with superior oxygen uptake. Nat Commun 9:1378
191. Altintas C, Erucar I, Keskin S (2018) High-throughput computational screening of the metal organic framework database for CH 4/H 2 separations. ACS Appl Mater Interf 10:3668–3679
192. Azar ANV, Velioglu S, Keskin S (2019) Large-scale computational screening of metal organic framework (MOF) membranes and MOF-based polymer membranes for H 2 /N 2 separations. ACS Sustain Chem Eng 7:9525–9536
193. Inokuma Y, Matsumura K, Yoshioka S, Fujita M (2017) Finding a new crystalline sponge from a crystallographic database. Chem – An Asian J 12:208–211
194. Zhang L, Chen Z, Su J, Li J (2019) Data mining new energy materials from structure databases. Renew Sust Energ Rev 107:554–567
195. Shi P-P, Tang Y-Y, Li P-F, Liao W-Q, Wang Z-X, Ye Q, Xiong R-G (2016) Symmetry breaking in molecular ferroelectrics. Chem Soc Rev 45:3811–3827
196. Cole JM, Kreiling S (2002) Exploiting structure/property relationships in organic non-linear optical materials: developing strategies to realize the potential of TCNQ derivatives. CrystEngComm 4:232–238
197. Phan H, Hrudka JJ, Igimbayeva D, Lawson Daku LM, Shatruk M (2017) A simple approach for predicting the spin state of homoleptic Fe(II) Tris-diimine complexes. J Am Chem Soc 139:6437–6447
198. Schober C, Reuter K, Oberhofer H (2016) Virtual screening for high carrier mobility in organic semiconductors. J Phys Chem Lett 7:3973–3977
199. Kunkel C, Schober C, Oberhofer H, Reuter K (2019) Knowledge discovery through chemical space networks: the case of organic electronics. J Mol Model 25:87
200. Cole JM, Low KS, Ozoe H, Stathi P, Kitamura C, Kurata H, Rudolf P, Kawase T (2014) Data mining with molecular design rules identifies new class of dyes for dye-sensitised solar cells. Phys Chem Chem Phys 16:26684–26690
201. Adalder TK, Dastidar P (2014) Crystal engineering approach toward selective formation of an asymmetric supramolecular synthon in primary ammonium monocarboxylate (PAM) salts and their gelation studies. Cryst Growth Des 14:2254–2262
202. Veits GK, Carter KK, Cox SJ, McNeil AJ (2016) Developing a gel-based sensor using crystal morphology prediction. J Am Chem Soc 138:12228–12233
203. Elton DC, Boukouvalas Z, Butrico MS, Fuge MD, Chung PW (2018) Applying machine learning techniques to predict the properties of energetic materials. Sci Rep 8:9059
204. Wicker JGPP, Cooper RI (2015) Will it crystallise? Predicting crystallinity of molecular materials. CrystEngComm 17:1927–1934
205. Directed Assembly Network (2020) Directed assembly themes and streams. http://directedassembly.org/themes-and-focus/
206. Grabowsky S, Genoni A, Bürgi H-B (2017) Quantum crystallography. Chem Sci 8:4159–4176
207. Lommerse JPM, Motherwell WDS, Ammon HL et al (2000) A test of crystal structure prediction of small organic molecules. Acta Crystallogr Sect B Struct Sci 56:697–714

208. Reilly AM, Cooper RI, Adjiman CS et al (2016) Report on the sixth blind test of organic crystal structure prediction methods. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 72:439–459
209. Cole JC, Groom CR, Read MG, Giangreco I, McCabe P, Reilly AM, Shields GP (2016) Generation of crystal structures using known crystal structures as analogues. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 72:530–541
210. Musil F, De S, Yang J, Campbell JE, Day GM, Ceriotti M (2018) Machine learning for the structure–energy–property landscapes of molecular crystals. Chem Sci 9:1289–1300
211. Bryant MJ, Black SN, Blade H, Docherty R, Maloney AGP, Taylor SC (2019) The CSD drug subset: the changing chemistry and crystallography of small molecule pharmaceuticals. J Pharm Sci:1–8
212. Fábián L (2009) Cambridge Structural Database analysis of molecular complementarity in cocrystals. Cryst Growth Des 9:1436–1443
213. Altomare A, Cuocci C, Giacovazzo C, Moliterni A, Rizzi R, Corriero N, Falcicchio A (2013) EXPO2013: a kit of tools for phasing crystal structures from powder data. J Appl Crystallogr 46:1231–1235
214. Feng ZJ, Dong C (2007) GEST: a program for structure determination from powder diffraction data using a genetic algorithm. J Appl Crystallogr 40:583–588
215. Toby BH, Von Dreele RB (2013) GSAS-II : the genesis of a modern open-source all purpose crystallography software package. J Appl Crystallogr 46:544–549
216. Coelho AA (2018) TOPAS and TOPAS-academic : an optimization program integrating computer algebra and crystallographic objects written in C++. J Appl Crystallogr 51:210–218
217. Schärfer C, Schulz-Gasch T, Ehrlich H-C, Guba W, Rarey M, Stahl M (2013) Torsion angle preferences in druglike chemical space: a comprehensive guide. J Med Chem 56:2016–2028
218. David WIF, Shankland K, van de Streek J, Pidcock E, Motherwell WDS, Cole JC, IUCr (2006) *DASH* : a program for crystal structure determination from powder diffraction data. J Appl Crystallogr 39:910–915
219. Kabova EA, Cole JC, Korb O, López-Ibáñez M, Williams AC, Shankland K (2017) Improved performance of crystal structure solution from powder diffraction data through parameter tuning of a simulated annealing algorithm. J Appl Crystallogr 50:1411–1420
220. Kabova EA, Cole JC, Korb O, Williams AC, Shankland K (2017) Improved crystal structure solution from powder diffraction data by the use of conformational information. J Appl Crystallogr 50:1421–1427
221. Cole JC, Kabova EA, Shankland K (2014) Utilizing organic and organometallic structural data in powder diffraction. Powder Diffract 29:S19–S30
222. Florence AJ, Bardin J, Johnston B, Shankland N, Griffin TAN, Shankland K (2009) Structure determination from powder data: mogul and CASTEP. Zeitschrift für Krist 2009:215–220
223. Shankland K, Spillman MJ, Kabova EA, Edgeley DS, Shankland N (2013) The principles underlying the use of powder diffraction data in solving pharmaceutical crystal structures. Acta Crystallogr Sect C Cryst Struct Commun 69:1251–1259
224. Florence AJ, Shankland N, Shankland K et al (2005) Solving molecular crystal structures from laboratory X-ray powder diffraction data with DASH : the state of the art and challenges. J Appl Crystallogr 38:249–259
225. Bruker AXS Inc (2012) SAINT. Madison, Wisconsin
226. Reymond J-L, Awale M (2012) Exploring chemical space for drug discovery using the chemical universe database. ACS Chem Neurosci 3:649–657
227. American Chemical Society (2020) Chemical abstracts service. https://www.cas.org/about/cas-content
228. Hey T, Tansley S, Tolle K (2009) The fourth paradigm: data-intensive scientific discovery. Microsoft Research

Struct Bond (2020) 185: 141–198
https://doi.org/10.1007/430_2020_70

Published online: 29 September 2020

Crystallography Under High Pressures

Stephen A. Moggach and Iain D. H. Oswald

Contents

Abstract This chapter highlights the area of crystallography of molecular systems under high-pressure conditions. It is an area of crystallography that has seen a rapid expansion over the last two decades. Advances in technology and data processing have facilitated the discovery of new materials, polymorphs and chemistries under

S. A. Moggach
School of Molecular Sciences, The University of Western Australia (M310), Perth, WA, Australia
e-mail: stephen.moggach@uwa.edu.au

I. D. H. Oswald (✉)
Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, UK
e-mail: iain.oswald@strath.ac.uk

extreme conditions. We discuss these advances using examples of organic and metal-organic materials as well as providing guidance to the pitfalls to be avoided conducting these studies.

Keywords Amino acids · Coordination polymers · Diamond anvil cell · High-pressure recrystallisation · Large volume press · Metal-organic framework · Molecular magnets · Pharmaceuticals · Pressure · Spin crossover

Abbreviations

DAC	Diamond anvil cell
MOF	Metal-organic framework
PTM	Pressure-transmitting medium
ZIF	Zinc imidazole framework

1 Introduction

High-pressure science is an exciting area of chemical crystallography that has developed significantly over the past 25 years. The strides that have been made in equipment and in data processing have permitted high-pressure crystallography to become an almost routine technique for the twenty-first century. The ability to probe materials under conditions four orders of magnitude more extreme than is practicably achievable by temperature has enabled the characterisation of many novel high-pressure polymorphs of materials that have shown little or no propensity for polymorphism at ambient pressure [1, 2]. This fact has promoted the use of pressure in the consciousness of scientists in many areas from pharmaceuticals to metal-organic framework materials. A key driver for many scientists is to connect the structure to the function of materials and in particular how the materials respond under their working environments. High-pressure crystallography can play a role in these studies by providing atomic level detail of the changes that occur in materials under these extreme conditions. From these studies, improvements in models can be made providing increasingly accurate descriptions of the processes involved. This has been proven to be the case for the metal-organic framework material, ZIF-8, where the conformational changes that occur at 1.5 GPa can be used to explain the behaviour of this material at high gas loadings [3]. Another area in which pressure can aid modelling is in the area of energetic materials. During operation these materials can experience high pressures as shockwaves pass through the materials prior to deflagration [4, 5]. Their characterisation, at the atomic level, can provide experimental evidence of changes, including polymorphic transitions that can be incorporated into models of deflagration, improving their accuracy. High pressure is also being used as a form of polymorph screening of pharmaceutical materials [6, 7]. Polymorph screening is a major part of the development of pharmaceutical

materials to ensure the stability of any solid form taken forwards to production. The potential financial impact of the appearance of an unknown and uncharacterised polymorph with different physicochemical properties is substantial, never mind the impact on public trust. Hence there has been a move towards using high pressure to survey different areas of thermodynamic space providing a more robust screen of potential drug products. So whilst high pressure may have been considered a niche area of research that is the reserve of earth scientists and physicists, it is certainly not the case now and it is thanks to the developments made by pioneering groups in the area of high-pressure crystallography many of whom will be referenced in this chapter. High-pressure crystallography is a powerful technique that can provide an alternative pathway to explore molecular materials and should be a consideration of any solid-state scientist moving forwards through the twenty-first century.

As a preface to the chapter, the work highlighted herein reflects work conducted by our respective groups placed in the wider context. We have divided the chapter into different sections so that readers can focus on particular areas of interest. Section 2 details the different methodologies that have been employed to investigate materials and the impact these have made on the systems under study. We have tried to draw on the studies of many groups working in high pressure, and where possible we have linked the examples in Sections 3 and 4 to the methodologies used where appropriate and where the method may have impacted on the observed results. This is not a complete review of the effects of high pressure on molecular materials, but we hope that it provides an overview of the types of work that have been conducted. There are a number of reviews of molecular materials at high pressure that we would encourage the reader to engage with on subjects such as amino acids [2, 8], energetic materials [5, 9], metal-organic frameworks, chemical reactions as well as the basics of the technique in general [10, 11].

2 High-Pressure Methodologies

2.1 Standard Methods

In this section we will discuss various methods that can be employed to investigate organic and metal-organic compounds under high-pressure conditions. The choice of method will vary depending on the science to be investigated as, in the experience of the authors, some methods can facilitate phase transitions more readily than others. The practicalities of loading diamond anvil cells (DACs) using liquids or single crystals will not be discussed here as they have been covered very well in other texts; however, we will indicate considerations that may need to be taken into account in certain circumstances. The authors would recommend the chapter ‘High Pressure Single-Crystal Techniques’ by Miletich et al. to provide a comprehensive review of the method [11].

2.2 *Loading Methods*

2.2.1 Low-Melting Compounds

Some of the early work in single-crystal high-pressure crystallography investigated the difference in the crystallisation behaviour of liquids under cooling or through compression. Low-melting compounds are still being investigated as they simplify the nature of interactions between molecules so that, for example, the effect of halogen bonding may be investigated more specifically [12, 13]. The preparation for the loading of low-melting compounds is no different to any other high-pressure loading and, in some ways, a little easier as there are fewer items requiring loading, i.e. ruby and liquid. There are a few considerations that need to be taken into account for a successful experiment such as the maximum pressure of the experiment to be conducted and the volatility of the liquid in question. For most one-off crystallisation experiments at lower pressures, gasket holes with a large diameter (250–300 μm) and greater depth (100–120 μm) permit a larger crystal to be grown. This is advantageous as it provides potentially stronger diffraction from the crystal. However, it does limit the maximum pressure at which the sample can be analysed, but, in most cases in the literature, a maximum of 1.5 GPa would be the norm as the melting curve generally increases with pressure thereby limiting the maximum pressure at which the liquid can be heat annealed. To load the cell, a ruby chip is placed in the sample chamber (or in the opposing diamond) and a drop of the liquid under study is placed over the hole ensuring that there are no significant bubbles present in the chamber; the ease of this process will depend on the volatility of the liquid. Through trial and error, for the best success and to allow for some evaporation, use the surface tension of the liquid to maximise the liquid on the gasket surface which also provides extra time to deal with any bubbles that may be present in the hole. Once the cell is sealed, there may be a small bubble present, but usually these are solubilised on compression.

Nucleation of products at pressure can be troublesome as liquids can form glasses on compression and the pressure increases that are used to initiate the crystallisation can be rapid; hence, kinetics can play a significant role. Nevertheless nucleation can occur through a number of different methods from compression of the liquid alone [14], pressure cycling (varying pressure up and down) [15], through use of liquid nitrogen to freeze the sample [16] or applying heat [17]. Each of these methods has successfully been used to nucleate samples, and one may need to explore all of these methods for a successful outcome. Further recent methods will be explored when we discuss quenching and recovery of samples (Sect. 2.3.2.1).

In general, the nucleation product is polycrystalline which can be analysed using X-ray powder diffraction or spectroscopic methods, but for those who wish to perform single-crystal diffraction, there is a requirement to anneal the polycrystalline material into one crystal or a reduced number of crystallites for analysis. Two main methods used are heat annealing and pressure annealing. Both methods require that the sample is close to the melting line of the compound which may not

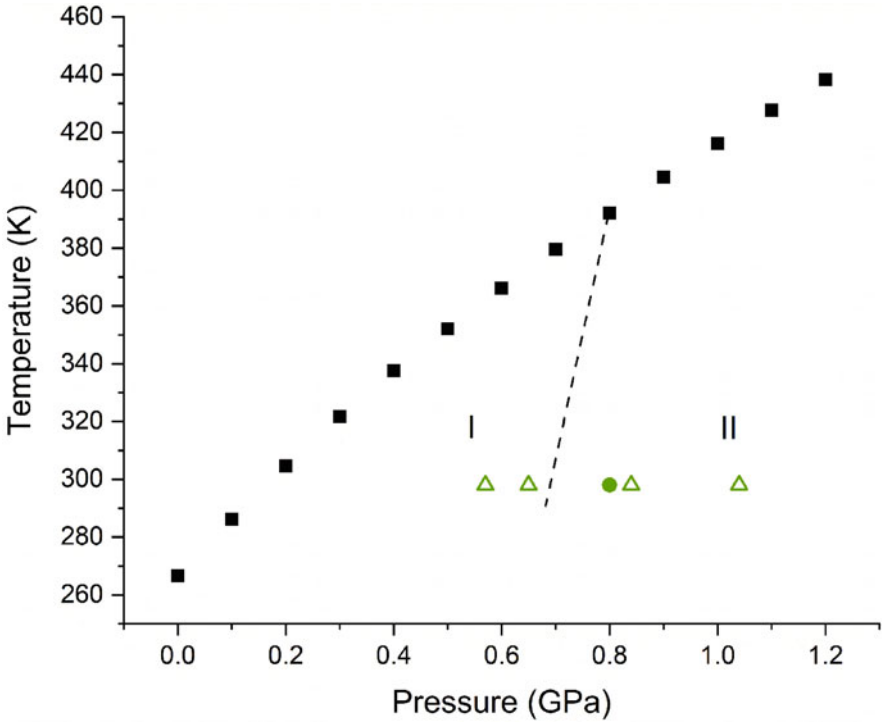


Fig. 1 The pressure/temperature phase diagram for aniline. The melting curve points (black squares) were taken from Bridgman [18], and the structure determinations (green data points) were taken from Funnell et al. [22]. The dotted line has been added by us and represents a tentative phase boundary between Forms I and II directed by the commentary in Funnell et al.

necessarily be close to the nucleation pressure. An extreme example of this is methanol where the solid was nucleated at 7 GPa but the crystal was finally grown at 4 GPa due to the inability to melt the sample at 7 GPa [15]. The melting line of the sample can be determined solely from observing the reduction in the crystallite size during isothermal pressure annealing, but other methods have used volumetric measurements to ascertain it. Work by Bridgman [18] and more recently Dziubek and Katrusiak [19, 20] have shown that not only can the melting line be observed but new high-pressure phases can be observed if there is a discontinuity in the pressure-volume plot; in fact, these measurements provided the basis for our study and discovery of a new polymorph of 2-methylphenol at 0.65 GPa [18, 21].

Once the melting point has been determined, this can be used to anneal the polycrystalline material into a single crystal through pressure cycling. The isolation of novel phases through isothermal pressure cycling will be limited due to the reliance on the phases being stable at a particular pressure at ambient temperature. To access novel forms of a low-melting compound, there may be a requirement to add heat to be able to access new forms. Figure 1 shows the pressure/temperature phase diagram for aniline. Aniline demonstrates a couple of interesting points of

discussion in relation to low-melting compounds: (1) the potential need for temperature to aid access to polymorphs and (2) the metastability of phases due to kinetics. Annotated on the pressure/temperature phase diagram are the experiments of Bridgman [18] (melting line, black squares) and Funnell et al. [22] (green data points) as well as a tentative phase boundary between Forms I and II (our addition). The solid green circle designates the single-crystal data for Form II from a crystal that was nucleated at 0.8 GPa and annealed with temperature to reach the melting line at ~390 K. On cooling, the crystal was stable (or sufficiently metastable) at ambient temperature for a dataset to be collected. In this case, a temperature increase was necessary to be able to access the new form. Had the crystal been grown around the melting line at ambient temperature (0.15 GPa), the known Form I would have been observed. This is quite common, and we have observed this with 3-chlorophenol and 3-fluorophenol where we determined the structures at very low pressures of 0.1 and 0.12 GPa [23]. In cases like these, it is worth applying further pressure until one is no longer able to reach the melting curve on heating. To access pressures beyond this, recrystallisation methods will be required (Sect. 2.2.3). On a practical note, the speed of melting and growth can be altered by how far away from the melting curve the experiment is being conducted; the closer to the melting line, the slower the crystal growth will be, and the growth can be accelerated by performing at higher pressures due to being deeper into the phase stability region. Small changes in the pressure can help in the crystal growth if, for example, many crystallites continue to nucleate during the growth stage.

The second point of discussion is the metastability of phases beyond phase boundaries. This behaviour is common where a large rearrangement of molecules is required between the two phases. From neutron powder diffraction measurements, Funnell et al. were able to isolate Form I of aniline at 0.65 GPa, which has allowed the tentative assignment of the phase boundary, and compress Form I to 1.04 GPa without any transformation taking place (hollow green triangles). From the structure of the two phases, they were able to explain this metastability through the large differences in structure which forms a large kinetic barrier to overcome. We have also observed this with 2-methylphenol where it crystallises in $P3_2$ at low pressures and can be compressed to 8 GPa without any indication of transition to the high-pressure monoclinic phase ($P2_1/c$) due to the substantial change in structure required [21]. This has also been noted in high-melting point compounds such as glycine [2, 8] and imidazole [24].

As a final note in this topic, in our experience of data processing of high-pressure materials, we have been able to deconvolute three to four individual diffraction patterns and merge the datasets to a satisfactory outcome [1, 25]. Our advice to those starting out in the area of crystal growth at high pressure would be to err in the side of caution and grow multiple crystals, collect data on them and try to process the datasets derived from these crystallites before attempting to grow one single crystal. The difficulty of nucleation means that the loss of a crystal can lead to further hours of nucleation experiments to obtain the polycrystalline material again. In addition, the multiple crystals can lead to a substantial increase in the data completeness after

merging datasets especially combined with collecting datasets with the DAC in different orientations, i.e. mounting a triangular DAC along another side.

2.2.2 Mixing Product with PTM

For systems that show no phase transformations up to a maximum pressure where they can be melted, there are other routes that can be explored. In the case of aniline [22] and other systems such as acetone [26] and ethanol [14], these were loaded and solidified from the pure liquid; hence, the experiments that can be performed are limited. These limitations are either due to the melting point becoming too high for successful melt and crystal growth or due to the compression of the sample becoming non-hydrostatic beyond a certain pressure. A method to circumvent some of these issues is to mix the liquid or gas of choice with another to increase the pressure at which the crystals can be grown or to ensure the hydrostaticity to higher pressures. Much of the inspiration for this work comes from the work on gases and gas clathrates where the gas of choice, e.g. nitrogen, is mixed with a second gas to act as a pressure-transmitting medium (PTM), e.g. helium [27–30]. The PTM allows the compound under study to dissolve rather than melt which can lower the barrier to interconversion. A number of studies have used this approach to facilitate the growth of new phases. We have investigated the high-pressure phase behaviour of acrylic and methacrylic acid using this technique due to the chemical reaction (polymerisation) that occurred on heating these systems at high pressure and high temperature [31–34]. Using a 50:50%v/v mixture with 4:1 methanol/ethanol, we were able to grow crystals of acrylic acid at much higher pressures than could be accessed through conventional routes (0.65 GPa acrylic acid; 1.5 GPa methacrylic acid). The dissolution of the compound under study in the PTM helped to lower the barrier to interconversion as the solvation energy helps to overcome the energy barrier [9]; this principle will be employed in Sect. 2.2.3. Other positive outcomes from adopting this method are that phase transformations can be accelerated at ambient temperature by use of the solution. Crystals of acrylic acid were observed to transform when the pressure exceeded the Form I-Form II phase boundary [33]. Katrusiak and co-workers have used this method to investigate o-xylene so that higher pressures could be achieved [35]. Using a neat liquid, the authors could only achieve a pressure of 0.31 GPa and still have the ability to melt the sample. By mixing o-xylene with methanol in a 1:4 and 1:9 ratio, they were able to increase the pressures to 1 and 3.5 GPa, respectively. This is a huge expansion of phase space that can be explored. In this case, Marciniak and Katrusiak did not observe any new polymorphs of o-xylene, but it does highlight the potential of the method for polymorph discovery.

2.2.3 Recrystallisation

In recent years the use of solutions to explore the solid-state forms has become more prevalent as it provides an orthogonal approach that can be used to explore materials that possess melting points that are too high to anneal with increased temperature. To explain the methodology, we will use a series of studies of molecular systems to demonstrate a large number of scenarios that we and other researchers in this field have observed.

Paracetamol

Prior to 2003, paracetamol was known to exist in two polymorphic forms: the thermodynamically stable monoclinic $P2_1/c$ phase [36] and the metastable orthorhombic *Pbca* phase [37]. The interest in paracetamol lies in the fact that the orthorhombic polymorph possesses better compaction properties compared with the stable monoclinic polymorph. The metastable form can be isolated easily from the melt and is stable for approximately 50 h during experiments to test sublimation [38] or 30 min in contact with solvent [39]. Paracetamol had been explored using high pressure via compression methods by the group of Boldyreva. Their paper demonstrated that through compression one could induce the transformation to the orthorhombic phase in a powder sample through pressure alone although the transition was incomplete and slow due to the large rearrangement of the molecules required between the two phases [40]. To aid the isolation of the orthorhombic form, Fabbiani et al. investigated the use of recrystallisation at high pressure [41]. By using solutions in this way, they utilised the solvation energy to help overcome the kinetic barrier to transformation. In their study they loaded a ca. 1 M methanol solution of paracetamol into the DAC and treated the solution in the same way as pure liquids had been investigated in prior studies through heat annealing. From this solution, a new methanol solvate was precipitated, and by repeated heat/cool cycles, a single crystal was formed that could be analysed by single-crystal X-ray diffraction. This was a significant and novel approach to materials preparation, and it opened the door to the analysis of high-melting compounds such as pharmaceutical products.

We have been employing these methods routinely and have found that a key difference from the manipulation of pure liquids is the reaction of the crystals and solution to heat. Crystals obtained from a pure liquid react almost instantaneously to the application of heat or on cooling; hence, the growth of the crystal can be controlled very easily. Generally, the response to heating or cooling of solutions is, naturally, much more delayed due to the dissolution process; hence, even more care is required when trying to isolate one crystallite during the annealing procedure. As with the pure liquid method, the pressure can be tuned to increase or decrease the rate of crystal growth, but there is still a lag present.

Further studies of paracetamol at high pressure have indicated that the solid form observed can be dependent on the solvent as well as the pressure at which the phase

was precipitated. We were able to isolate the orthorhombic phase using the pressure-precipitation methodology at a range of pressures from 0.2 GPa (acetone) to 1.1 GPa (ethanol) in a DAC without any indication of further phases in these solvents [42, 43]. However, from a 0.06 M aqueous solution at 1.1 GPa, a dihydrate can be formed, but, given the pressure that it was formed, there may be a drive for inclusion of water into the structure [42]. Other observations of this kind have been observed for piracetam, where the precipitation occurred with the solidification of ice [44], and the crystallisation of γ-aminobutyric acid [45]. The relative stability of the pure paracetamol polymorphs at high pressure was confirmed by high-pressure differential scanning calorimetry studies which identified that the orthorhombic phase is the most thermodynamically stable form at high pressure between pressures of 0.29 and 0.45 GPa (maximum pressure of the calorimeter) [46]. Even scale-up to larger volumes was possible through the use of large volume press and quench cooling [43].

Recently we have developed the method for performing anti-solvent addition at high pressure using the large volume press [47]. We investigated this methodology to aid the crystallisation of compounds at high pressure as well as to potentially stabilise the recovery of high-pressure solid forms. We chose paracetamol as an exemplar material as the literature on its high-pressure behaviour was extensive. In this study we chose a mixed solvent system that showed a rapid decrease in solubility with addition of anti-solvent (64% w/w aqueous methanol). One of the challenges with this method is that the large volume press is not transparent so we complemented these studies with DAC work. We were able to show that precipitation can occur and that the high-pressure orthorhombic form could be recovered to ambient pressure. During this investigation, using the DAC, we observed a further methanol solvate that is structurally similar to the known phase but possesses a $Z' = 3$. Due to the small changes in the structure, we surmised that the change in the solvent used may have had a subtle effect on the structure of the crystallised form that manifests itself as a change in the crystal structure. Figure 2 depicts the phase behaviour of paracetamol through recrystallisation methods under the range of conditions.

Piracetam

Piracetam is a fantastic example of how the concentration and pressure can be combined to access multiple different solid-state forms (Fig. 3). Fabbiani et al. explored the use of different concentrations of solution in different solvents as well as simple compression techniques to successfully identify four different solid modifications of piracetam (Forms III–V and a dihydrate) [44, 48]. The ‘simple’ compression technique applied here involved the annealing of the powder at low pressure in 2-propanol to bypass the issues of cutting the crystal to the correct size before the application of pressure. In this system, there is a general observation for the precipitation experiments that at higher concentrations of solution, Form IV could be accessed (1.6 M methanolic solution or 6 M aq. solution) whilst at lower

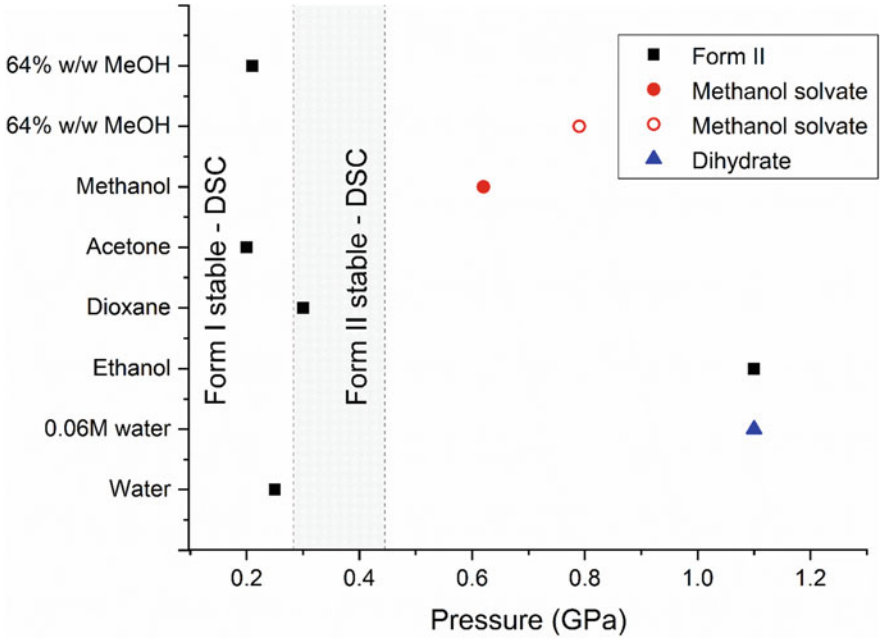


Fig. 2 The solid-state forms of paracetamol obtained from the recrystallisation at high-pressure methodology. The shaded region designates where it has been shown that paracetamol form II is stable through differential scanning calorimetry of the pure compound [46]. Filled black squares are Form II; filled red circle is methanol solvate from Ref. [41]; open circle is methanol solvate from Ref. [47]; filled blue triangle is dihydrate from Ref. [42]

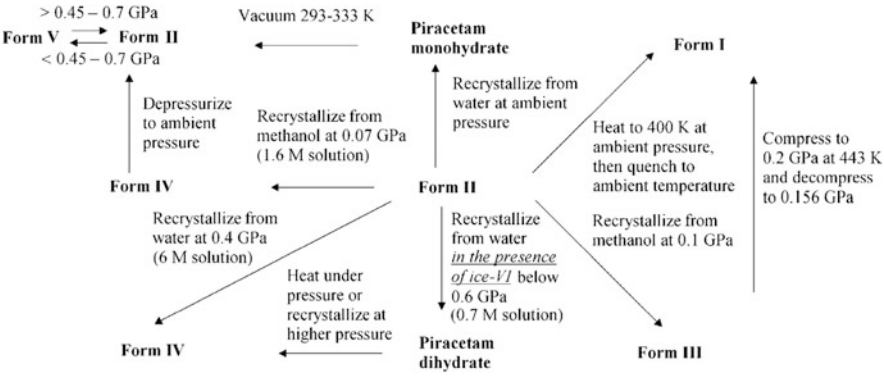


Fig. 3 A schematic representing the extensive solid-state behaviour of piracetam at high pressure. Forms III–V are high-pressure polymorphs of piracetam, and the dihydrate can only be accessed at high pressure through precipitation of the solvent aqueous medium. Reprinted with permission from Cryst. Growth Des., 2007, 7, 1115–1124. Copyright (2007) American Chemical Society [44]

concentrations the Form III or the dihydrate could be isolated. The latter observation of the dihydrate is interesting as the lower concentration enabled precipitation at higher pressure to isolate the hydrated form. They had to crystallise the solvent as ice-VI at 1.6 GPa before decompression to remove ice and heat annealing of the piracetam powder. Paracetamol dihydrate and piracetam dihydrate can both be formed through the crystallisation of the solvent before decompression to reveal a powder that can be annealed. In each of these cases, ice-VI could be acting as a nucleator and templating the hydrate. This is supported to some extent by the crystallisation of Form IV using more concentrated aqueous solution that precipitates at lower pressures. Alternatively, in a study of γ-aminobutyric acid, Fabbiani and co-workers were able to provide computational evidence to support the stabilisation of the hydrate over the pure compound; hence, the hydrate is the favoured product at pressure [45]. This idea is certainly worth exploring given γ-aminobutyric acid hydrate is recoverable to ambient pressure and hydration may be a route to provide stable forms of pharmaceutical products. The tuning of experimental conditions in this way demonstrates the complexity of the high-pressure recrystallisation experiment and the potential to realise novel polymorphs and solvates of materials under these conditions.

2.2.4 Compression

The previous topics have described how the compound or solvent can be manipulated to isolate new solid forms of materials at fairly low pressures. They have been limited to lower pressures due to the melting point of the compound of study or the freezing pressures of the solvent in which the solute has been dissolved. Compression studies have been the standard method to investigate materials at high pressure. These studies can provide evidence of how intermolecular interactions evolve over a pressure range or whether the physical properties of a material may change. The changes to the experimental set-up for these experiments are limited to the geometry of the sample environment. The gasket hole size and thickness of gasket are altered depending on the pressures that need to be achieved. The choice of PTM is also critical to ensure hydrostaticity of compression, but it can play a large role in how the compound of study reacts to pressure changes. In this section, we would like to highlight two recent studies that have demonstrated that compression studies are not routine and that changing PTM can alter the behaviour of molecular materials at pressure. This will be explored further in Sect. 4.5 highlighting metal-organic materials at high pressure.

Collings and Hanfland have recently investigated 4-hydroxycyanobenzene in relation to charge transport properties for semiconductors and how pressure can play a role in changing the intermolecular interactions and hence the properties of the material [49]. In their study they noticed that the PTM was playing a role in the formation of new host-guest phases of 4-hydroxycyanobenzene. The insertion of PTMs into the crystal structure is quite unusual for organic structures that do not possess a channel structure, but it is well-known in metal-organic framework

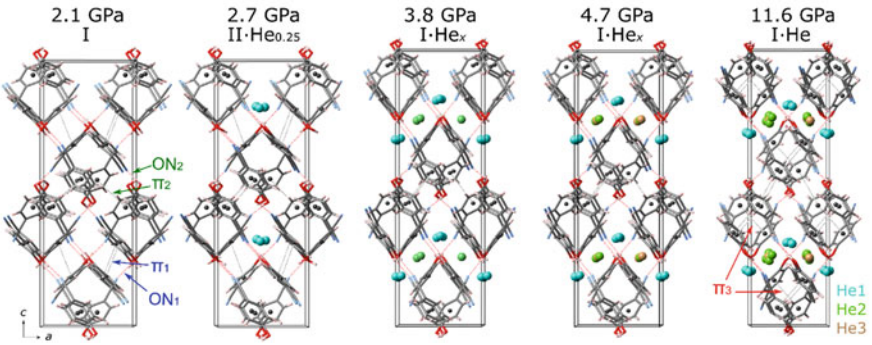


Fig. 4 The introduction of helium into the crystal structure of 4-hydroxycyanobenzene on application of pressure. Figure taken from Molecules 2019, 24(9), 1759, DOI: (https://doi.org/10.3390/molecules24091759) [49]

materials [50]. During the compression using helium as the PTM, they noted that above 2.5 GPa, helium inserted into the structure between the hydrogen-bonded columns in regions that could not be classed as voids (Fig. 4). The insertion of helium impacted on the symmetry of the structure, and through the use of single-crystal diffraction, they were able to ascertain the sequential loading of helium into the structure at particular sites, a common observation in framework materials [50]. The helium-doped sample remains stable to 26 GPa but loses helium on decompression to 2.5 GPa. The authors turned to using neon as the PTM for the second experiment to prevent the inclusion into the structure. Through compression alone they observed two new phases at 4 and 5.8 GPa where Form II is an intermediate phase and only a portion of molecules in the structure transition to the new configuration. This study demonstrates that even in organic molecular systems the choice of PTM is crucial to the observations that we make, particularly when gases or smaller PTMs are used.

The choice of PTM can be important even with liquids as was demonstrated by Eikeland et al. who chose to compress the clathrate structure, hydroquinone/formic acid, in two studies using 1: 1 pentane/isopentane and silicone oil; a third was conducted in paratone-N oil, but no data were provided [51]. The behaviour of the clathrate structure is markedly different with a phase transition occurring at ~4 GPa using the pentane mixture, whilst the transition is hindered when using the silicone oil or paratone-N oil. The authors attribute the difference in behaviour due to the solid nature of the silicone oil at pressures above 1 GPa whilst the pentanes mixture is still liquid allowing a physical change in the bulk of the crystal. In the metal-organic framework UiO-abdc, changing the PTM can have significant effects on the compressibility, with the bulk moduli varying by an order of magnitude, with a greater than 10% difference in compressibility (see Sect. 4.5). In the case of the clathrate hydroquinone/formic acid, arguments of solubility were highlighted in this case with respect to traces of water in the pentanes mixture; however, a more likely reason is that the clathrate structure is soluble in the pentane/isopentane mixture, but

this was not considered. The structure itself demonstrates a 4% compression over the phase transition which together with the stabilisation of the host interactions provides the driving force for the phase transition to occur. The nature of host-guest effects in hydroquinone was explored further using methanol and acetonitrile as the guest molecules with reference to the unsolvated β-form of hydroquinone [52]. In this case, the solvents play a large role in stabilising the geometry of the host structure much like the helium PTM did in the previous example. The unsolvated form is unstable on compression at fairly low pressures (0.32(5) GPa) where it collapses into the α-form over a 10-h period. Supercompression of the phase to 1.2 (5) GPa was achieved for the limited time of a single-crystal X-ray diffraction data collection before it collapsed into a polycrystalline product. Kinetics obviously played its part as different crystals were used to construct the compression dataset using silicone oil albeit the transformation was also observed in pentane/isopentane mixture. The exploration of the host-guest structure with methanol or acetonitrile in the cavity proved that the ambient pressure structure was stable to much higher pressures using pentane/isopentane as the PTM (greater than tenfold increase in stability). In each of these cases, the structure undergoes a phase transition to new high-pressure phases at 6.2(1) and 4.0(1) GPa for methanol and acetonitrile, respectively. The lower pressure of transition is due to the size of acetonitrile and its alignment in the cavity. In the methanol solvate, it is the host that changes geometry resulting in the phase transition [52]. The bulk modulus of the low-pressure phases is relatively similar (methanol, 8.3(12); acetonitrile, 8.5(3); formic acid, 13.6(4) GPa [51]), and the high-pressure phases of acetonitrile and formic acid clathrates also have a similar bulk modulus at the onset pressure (acetonitrile, 42(1), and formic acid, 38(2) GPa). Further work in the area of host-guest complexes will be discussed in more detail in Sect. 4 where we focus on the effects of pressure on metal-organic framework materials.

In summary, for many materials the choice of PTM for compression may be limited to the solubility of the compound in the medium, but these examples highlight issues that can occur especially given the drive towards investigating framework materials at high pressure.

2.3 Developments

In this section we will explore the developments that have been made in the area of high pressure and those that will have a positive impact on science going forward.

2.3.1 Merrill-Bassett Diamond Anvil Cell (DAC)

The Merrill-Bassett design of DAC was a pioneering and highly impactful addition to the area of high-pressure science [53]. The ability to perform single-crystal diffraction experiments using a standard goniometer has enabled high-pressure

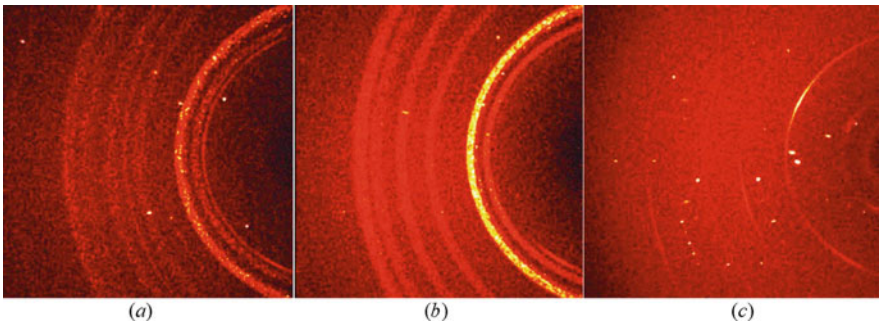


Fig. 5 CCD images of (**a**) a Be backing-seat cell collected using a Mo X-ray source for 60 s, (**b**) a Be backing-seat cell collected with synchrotron radiation ($\lambda = 0.6755$ Å) for 1 s and (**c**) a WC backing-seat cell collected using a Mo X-ray source for 60 s. Note that the Be powder lines are much more textured ('spotty') in (**b**) than in (**a**). Reprinted with permission from J. Appl. Cryst., 2008, 41, 249–251, DOI: https://doi.org/10.1107/S0021889808000514 [55]

science to be conducted routinely in a laboratory environment. Its simple design and ease of use has resulted in it becoming the workhorse of many high-pressure laboratories. The original design of the DAC used beryllium discs as backing plates for the diamonds to sit on. The beryllium plates provided a mechanically strong material whilst being relatively transparent to X-rays due to the low electron count of beryllium. The backing discs produced diffraction rings themselves which increased the background on the detector. This impacted on the data quality, but the development of the equipment enabled the collection of data for materials at high pressure on a laboratory instrument, a huge advancement in the technique as a whole. In 2004, Boehler and De Hantsetters developed a new design of diamonds and tungsten carbide backing seat that was strong enough to support the diamonds during compression but provided a large enough opening angle that the diffraction data were not absorbed by the WC seats [54]. Their design, which has since been incorporated into a Merrill-Bassett-type cell [55], is a canonical design where the diamonds are countersunk into the WC seat. The geometry of the set-up provides the support to the diamonds whilst maintaining the opening angle for observation of X-rays. Figure 5 shows the improvement in the diffraction by employing this new design. Other groups have used other methods such as diamond plates [56, 57] or beryllium-free cells [58] to reduce the background and analyse more complex samples [59].

Much of the work described herein will be based around the characterisation of materials through X-ray diffraction methods; however, there have been some recent developments in the area of neutron diffraction that will be applicable to readers. There are a number of designs of neutron cell that have been developed recently by Grzechnik et al. [60], to work with 'hot' neutrons ($\lambda = 1$ Å) at HEiDi at the Heinz Maier-Leibnitz Zentrum (MLZ), and Haberl et al. [61] to work at various instruments at Oak Ridge Laboratory's Spallation Neutron Source; however, these cells are solely designed for neutron diffraction. The methods employed by Binns and

co-workers enable the use of Merrill-Basset-type DACs with standard-sized samples on KOALA single-crystal diffractometer at ANSTO using Laue diffraction [62]. These experimental parameters open up the possibility of probing materials using neutron and X-ray techniques with the complementary information that they provide. To achieve this, a miniaturised DAC had to be made from beryllium-copper (BERYLCO-25) with necessary changes in design to support the pressure; however, standard Boehler-Almax diamonds were used. The beryllium-copper design allowed the cell to be fit and cooled in the cryostat on KOALA. Using hexamethylenetetramine and L-arginine dihydrate, the authors were able to demonstrate a high resolution and data completeness in the DAC (similar to one outside of the cell), and this was due to the observation of reflections through the cell body that provided enough data for anisotropic refinement at low temperatures. There are complications to the method such as centring and the requirement for larger-sized sample crystals ($0.15 \times 0.20 \times 0.30$ mm crystal was used in the study) lowering the maximum achievable pressure to ~5 GPa; however, this method provides a basis from which organic materials at high pressure can be explored using neutron diffraction.

2.3.2 Large Volume Presses

High pressure has been seen as a niche subject area for many years, but the strides that have been made into the exploration of pharmaceutical materials have moved this discussion onto industrial relevance of high pressure. Whilst much of the crystallographic work has been performed using DACs and Paris-Edinburgh presses, there have been developments in large volume equipment to investigate the formation and quenching of high-pressure phases for characterisation at ambient pressure.

Bridgman was the first to use large volume apparatus to investigate materials under high pressure using a hydraulic press [18]. His efforts to design the cell that would be able to make the measurements were not without its own hazards as he noted in his paper 'Finally, after six explosions, the attempt to use this form of apparatus was entirely given up . . .'. Despite these setbacks he was able to build a press that withstood the demands of pressure, and he began to build a legacy of pressure measurements that is second to none in the field. His press was made from two cylinders (upper and lower) connected by a heavy piece of tubing. The sample was housed in the lower cylinder, whilst the pressure was applied to the upper cylinder via a hydraulic press. The cylinder assembly was fitted with a micrometre so that the distance of compression could be measured and coupled with the diameter of the cylinder to assess the change in the volume of the sample; a manganin coil was used to assess the pressure of the cell through a change in its resistance. During the experiment the upper cylinder that contained the manganin coil was maintained at 35°C so that he did not need to worry about temperature effects on the resistance, whilst the sample was housed in a separate thermostat to alter the temperature whilst under load. To measure the melting points of the compounds, pressure was applied to the sample beyond the freezing point of the sample before taking readings at various pressures on decompression. To ensure that he characterised the melting line

well, he made sure he had 3–4 measurements either side of the melting line as well as two at the equilibrium where there was part melt of the sample. This procedure was repeated at various pressures. In his paper in 1914, he explored the melting curves of 11 substances ranging from simple elements (potassium) to more complex molecular systems such as aniline and o-cresol that has laid the groundwork for more recent evaluations of these systems through diffraction methods [21, 22, 63]. In the 1940s he continued to explore hundreds of more compounds in a series of papers after the development of new pieces of apparatus.

More recent developments in the area have been made by Dziubek and Katrusiak and our own work together with the Kamenev group and developments at ISIS. The work of Dziubek and Katrusiak developed apparatus capable of investigating the compression of materials on a large scale [19]. Their simple and effective piston design allows for the change in volume to be measured as a function of the applied load. In this set-up there is no control over sample temperature so increases in heat through adiabatic routes have to be minimised through small pressure increments. The body of the cell is large enough that the heat can be dissipated relatively easily. The effectiveness of the press was demonstrated using a number of simple organic compounds where (1) the phase transition was easily observed, e.g. chloroform, 3-aminopropan-1-ol and 1-methyl-benzoate, and (2) the phase transition was supressed by the large molecular rearrangement required from Phase I to Phase II, e.g. imidazole, akin to that observed for aniline [22] and 2-methylphenol [21].

In collaboration with the Kamenev group at the University of Edinburgh, we have built on previous work of the Pulham group and developed the use of our own large volume press (Fig. 6). Pulham and co-workers investigated the recovery of high-pressure forms to ambient pressure in quantities sufficient for characterisation at ambient pressure. The recovery or quenching of materials is particularly significant for fine chemicals, including pharmaceuticals, due to the potential for polymorphism in these systems. Polymorph screening is a vital technique to solid-form discovery, and it is conducted at ambient pressure by variation in solvent system, method of crystallisation, grinding, temperature, etc. By use of a large volume press for recovery, the program of discovery becomes far more extensive, and confidence in phase stability can be increased. Pulham and co-workers were able to use a large volume press, previously used to enhance Diels-Alder cycloaddition reactions for poorly reacting dienophiles [64], to produce the metastable orthorhombic form of paracetamol and recover it in gram quantities to ambient pressure through cooling for analysis using neutron diffraction [43]. We have continued to investigate the quenching of materials using this press and have successfully demonstrated that it can be achieved using exemplar systems of glycolide [65] and *p*-aminobenzoic acid [66].

To develop the methodology further, we have recently investigated the use of anti-solvent addition at high pressure using the large volume press. Nucleation is one of the major challenges when investigating liquids or solutions under high-pressure conditions; hence, methodologies that can improve and reproducibly initiate nucleation are a significant development. Anti-solvent addition is common practice in crystallisation methodologies at ambient pressure, but our ability to do this in a DAC

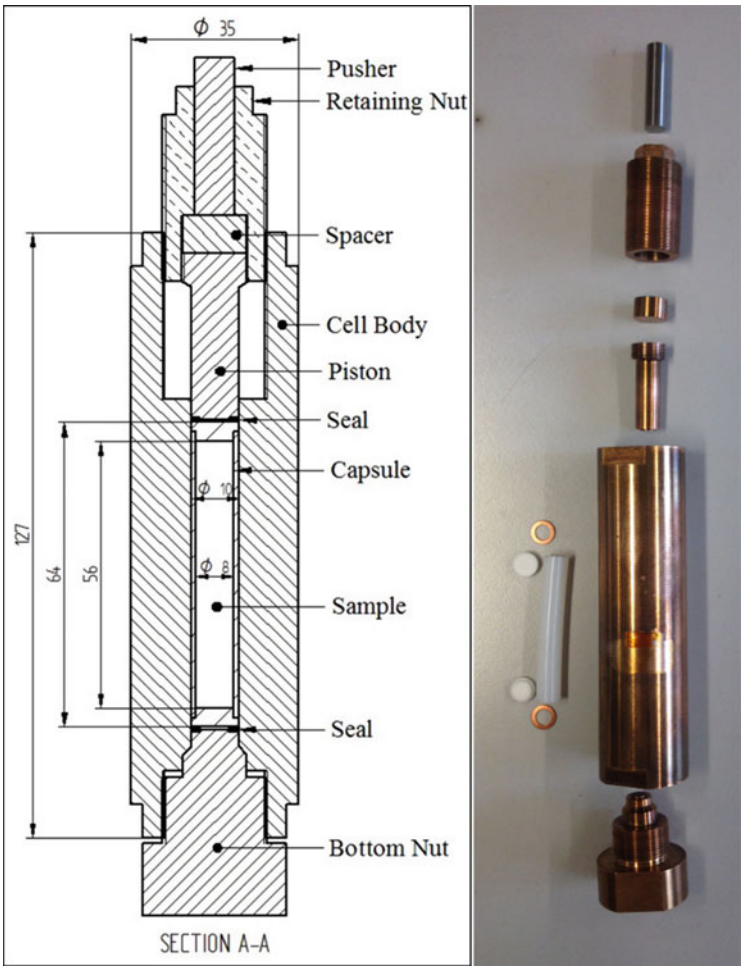


Fig. 6 The schematic and photograph of the piston-cylinder large volume press developed by Kamenev (University of Edinburgh) that has been used by our group to recover high-pressure polymorphs to ambient pressure

is restricted due to the sample volume. The large volume press, however, is ideal to address this challenge due to the size of the sample volume. Our design of cell is not transparent; hence, we needed to develop a methodology that would be independent of visual confirmation. By measuring the compression of the cell through the travel of the piston, and knowing the pressure in the cell, we were able to design glass inserts of the correct geometry to enable their fracture at a specific pressure point. We demonstrated that through the addition of the anti-solvent, we were able to recover the orthorhombic form of paracetamol to ambient pressure [47]. This was the first demonstration of this type of crystallisation process in a reproducible manner. There

are challenges to be overcome, such as solution-mediated transformations; however, the potential for the stabilisation of high-pressure forms in larger quantities is there.

Quenching of High-Pressure Forms

Whilst the identification of new high-pressure forms has been very successful, there have only been a few forms that have been quenched to ambient pressure. The observation of the recovery of materials from high pressure is likely to be underestimated as it has not been the primary goal of many pressure studies and hence may not have been recorded. The ability to quench novel forms to ambient pressure opens up the possibility for solid-form discovery and use as viable forms for delivery of pharmaceutical materials. Routes to manufacture through use of the large volume press and subsequent seeding are viable.

Fabbiani et al. investigated the neurotransmitter γ-aminobutyric acid (GABA) using pressure-induced recrystallisation techniques using a variety of concentrations from 6 to 12 M aqueous solutions as well as aqueous methanol solutions (4 M) [45]. In all cases, recrystallisation and crystal growth through temperature annealing formed a monohydrate of the GABA at 0.4 GPa which was the sole route to this novel hydrate. The enthalpy of hydration of the monohydrate was lower at all pressures and is the drive for its formation. Quenching GABA hydrate to ambient pressure was possible due to the similar enthalpies of hydration between the monohydrate and its constituent parts. In this case the monohydrate could be used as a seed in ambient pressure crystallisations. 5,6-Dimethylbenzimidazole shows similar behaviour where the hemihydrate is only accessed at high pressure [67]. The hemihydrate displays, unusually, a smaller molecular volume compared to the anhydrous form and can be recovered to ambient pressure and retained for several months. The thermodynamic drivers for the change were not investigated; however, the large volume change between the forms will contribute through the pV term of the free energy.

The most recent example from the pharmaceutical field is that of galunisertib [7]. Galunisertib has been investigated as a potential treatment for advanced metastatic malignancies and as part of that process was subject to solid-form screening protocols. It has a vast solid-form landscape that includes 50 solvated forms and 9 polymorphic forms. As part of this study, Bhardwaj et al. used crystal structure prediction to reveal hundreds of potential structures that were of high density and more importantly more thermodynamically stable than the known polymorphs. For these reasons high-pressure techniques were employed to explore this system. After a considerable search of the pressure phase space, they successfully elucidated the structure of a tenth polymorph (Form X, from a melt at high pressure; 0.4 GPa) that was shown to be recoverable to ambient pressure. Attempts were made to seed solutions; however, the small sample size restricted the ability to extend the experimental conditions.

Whilst recovery of materials from high pressure in a DAC is a step in the right direction, scaling up of the activity is where the major developments in this area will

lie. This has been demonstrated with the ε-phase of RDX [68], magnesium sulphate pentahydrate [69] and glycolide [70]. The epsilon phase of RDX is a high-temperature high-pressure polymorph. Miller et al. used the Paris-Edinburgh press to compress and heat RDX to 4.3 GPa and 448 K where they were able to identify a new phase through neutron diffraction. By quickly quenching the sample to ambient temperature, they were able to retain the phase and monitor the phase behaviour as a function of decompression to 0.75 GPa. Below this pressure, the polymorph was not stable and converted, but through quenching the sample to 150 K, they were able to recover it to ambient pressure. The phase persisted to 230 K before reverting to the stable α-phase.

More bespoke piston-cylinder equipment has been developed by Wang et al. that can be used to 2 GPa over a temperature range of 80–300 K in a laboratory setting [69]. This development has enabled the isolation and quenching of the high-pressure pentahydrate of magnesium sulphate to ambient pressure for characterisation using X-ray powder diffraction. Again, a low-temperature device (PheniX-FL) [71] was necessary to trap the high-pressure phase, but this is another example of a different technological solution to high-pressure recovery.

There are several other systems, such as 3-hydroxy-4,5-dimethyl-1-phenylpyridazin-6-one [72], cinchomeronic acid [73] and mefenamic acid [74], that have high-pressure phases that can be quenched to ambient pressure, but the longevity of the phase at ambient pressure was either short or not explored to any extent. These instances of quenchable phases are limited at present; however, as the technologies advance to aid stabilisation, this number will increase and the applicability to industrial processes will become more apparent.

3 Organic Materials Under Pressure

This section will provide an overview of the types of purely organic materials that have been investigated at high pressure. It is split into the various molecular types that have been our areas of interest such as alcohols, halogenated compounds, amino acids and pharmaceutically relevant materials. These have been chosen as the examples reflect some of the challenges that have been overcome by the use of techniques highlighted in Sect. 2.

3.1 Alcohols

Some of the simplest systems to be investigated have been the alcohols. The relative simplicity of the molecular structure and limited hydrogen bonding capability have provided an ideal set of systems to explore the interplay between hydrogen bonding and packing forces. In the 1990s there was an extensive body of work that investigated the solidification of liquids at low temperatures due to the development of

low-temperature devices [75]. To dovetail with these experimental efforts, in 1994 Brock and Duncan used the Cambridge Structural Database (CSD) to investigate the unusual distribution of space groups in which monoalcohols crystallised [76]. They were able to identify common structural motifs, e.g. crystallisation in $Z' > 1$ in low-symmetry space groups or $Z' = 1$ in high-symmetry space groups. This departure from common packing motifs was attributed to the size of the R-groups attached to the alcohol moiety. Even within the dataset, they observed different behaviour depending on the relative size of the R-group. For small R-groups, there was propensity to crystallise around twofold chains, but as the size of the alkyl group increased, higher fold chains, helices or discrete dimers were observed. Due to the effects of group size and the role of packing efficiencies in these systems, high pressure has been explored as a method to tune the molecular packing [14–16, 23, 77–80].

Initial studies on methanol [15], ethanol [14], phenol [81] and cyclobutanol [77] confirmed the hypothesis that by applying pressure to these systems, the packing of the molecules changed from being that of bulky alcohols to smaller alcohols: i.e. threefold axes changed to be linear twofold chains in the case of cyclobutanol; phenol is reduced from a pseudo threefold axis with $Z' = 3$ to a 2_1-screw with $Z' = 1$ (Fig. 7). 2-Butanol crystallises around a 2_1-screw axis at 2.14 GPa [82] which fits the hypothesis, but unfortunately it has not been explored at low temperature to confirm if it has the packing of a 'bulky' R-group. Other alcohols, such as 2-chlorophenol and 4-fluorophenol, have also followed a similar pattern, but it was apparent that this is not universal amongst all alcohols [80].

The change from cyclobutanol to cyclopentanol [78] sees a change in the behaviour as the latter adopts a flattened fourfold chain at 1.5 GPa that is attributed to a bulky alcohol rather than interaction as a twofold chain. However, the low-temperature behaviour of cyclopentanol is rather interesting as it exhibits four low-temperature phases that have yet to be fully characterised. The highest-temperature phases are orientationally disordered in hexagonal symmetry; hence, from the perspective of Brock and Duncan's rules, a change to a structure in which the molecules adopt a static position will be lessening the 'bulk' of the R-group; hence, one could say that pressure has altered the behaviour. Isopropyl alcohol [16] crystallises with unusual eight-membered hydrogen-bonded rings at 1.1 GPa with $Z' = 4$, and t-butanol [79] does not show any change in phase with respect to pressure to 0.85 GPa.

Building on the work on the alcohols, we investigated halogen-containing phenols to extend the size of the R-group and understand whether position of substitution would affect the outcome of the crystallisation [21, 23, 80]. At the time there was an increase in activity around organic fluorine which provided an interesting subtext to our research. Using the high-pressure recrystallisation techniques in combination with low-temperature studies and crystal structure prediction methods, we explored a range of monoalcohols, e.g. 2-, 3- and 4-fluorophenol and 2-, 3- and 4-chlorophenol. The behaviour varied markedly over the series of compounds and did not necessarily fit the observations that had been made previously regarding the change in packing at high pressure. 2-Chloro- and 4-fluorophenol showed behaviour

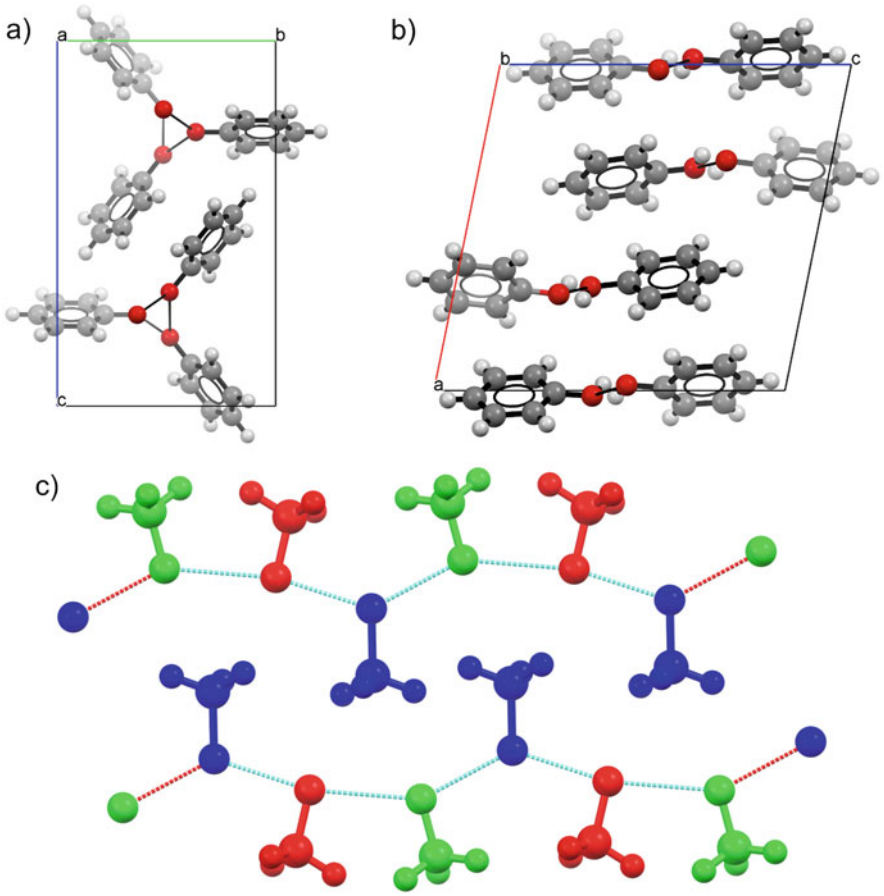


Fig. 7 (**a**) A typical hydrogen-bonded arrangement for molecules with 'bulky' R-groups as displayed by phenol at ambient pressure. Multiple molecules mimic a high-symmetry screw-axis (this case) or crystallisation of one molecule around a high-symmetry screw-axis (2-chlorophenol); (**b**) the high-pressure polymorph of phenol crystallised with one molecule in the asymmetric unit around a 2_1-screw axis, typical of the crystallisation of a 'thin' R-group; (**c**) the crystallisation of methanol at high pressure where it displays another type of packing where there is a 2-1-2-1 arrangement

that was reflective of phenol where the molecules are packed as linear chains at pressure [80]. However, 3-chloro- and 3-fluorophenol crystallised in the same polymorph at low temperature and at high pressure. Of the two polymorphs known for 4-chlorophenol (solid at 293 K), we found that Form II was more stable at high pressure and transformed to Form I on decompression, but in each case the structures possess packing attributed to a 'bulky' group. 2-Fluorophenol shows very interesting behaviour at high pressure. It crystallises in a new polymorph at 0.36 GPa compared with low temperature, but the molecules were arranged in a fourfold helix

akin to the 'bulky' R-group. During crystal growth the crystal fractured on cooling so that structure refinement was not possible. The data had to be collected at 403 K so that the crystal remained intact for the duration of the collection. It was speculated that the fracture of the crystal was due to an increase in the disorder on cooling [23]. Further work investigated the co-crystallisation of 2-methylphenol and 2-chlorophenol due to their isostructurality and demonstrated that both transform to the linear chain behaviour at high pressure. In the case of 2-chlorophenol, a new polymorph was discovered at even higher pressure than previous studies accessed through a failed co-crystallisation attempt with 2-methylphenol at high pressure [21]. More recently, Barnett and Allen have investigated the crystallisation of trifluoroethanol and observed two different phases they are composed of twofold chains. Interestingly, in this example, chemical pressure (substitution of hydrogen for fluorine) has induced a change in the packing of the molecules at low temperature to be similar to the high-pressure form of ethanol, whilst the high-pressure polymorph hydrogen bonds in a similar manner, but the packing of the molecules is more strained.

So there is a wide variety of behaviour that the alcohol systems exhibit. Nine out of 16 molecules studied so far alter their structure from a packing arrangement for 'bulky' R-groups to a packing for 'thin' R-groups with the application of pressure. Hydrogen bonds lengthen to enable the more efficient packing of these R-groups as the volume is reduced. The fact that methanol crystallises into another linear arrangement to promote packing that does not conform to ideal twofold symmetry (Fig. 7) does suggest that we may be able to take these systems and alter their packing even further by achieving higher pressures. The application of the mixed PTM techniques in Sect. 2 may be a route that would aid the discovery of new crystal packing of simple alcohols.

3.2 Halogenated Compounds

Halogen bonding is an area of expanding interest over the last few years. The manipulation of halogen bonding interactions is a critical tool in the area of crystal engineering where new molecular complexes are formed via this interaction. Readers are directed to an extensive review of the area by Cavallo et al. that highlights many of the advancements in the area [83]. We will provide an overview of how these interactions may be altered or used to great effect to promote chemical reactions.

3.2.1 Low-Melting Examples

There have been several studies that have investigated van der Waals halogenated solids to elucidate how halogen bonds are altered as a function of pressure in a similar manner to the studies of alcohols above. Some of the simplest molecules

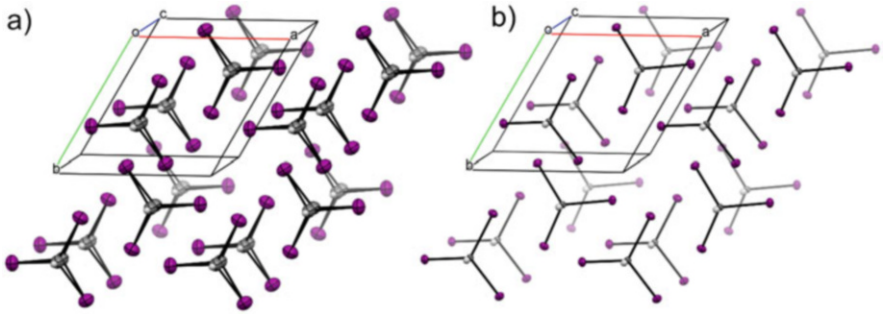


Fig. 8 (**a**) Crystal structure of iodoform in the ambient pressure $P6_3/m$ phase showing the disorder present due to the mirror symmetry perpendicular to the screw-axis, (**b**) the high-pressure phase of iodoform ($P6_3$) showing the ordered polar nature of the structure

investigated are chloroform and its bromo- and iodo-analogues. At high pressure all three systems are isostructural with one another crystallising in $P6_3$/m with the only difference being the pressure at which the high-pressure polymorph is observed. β-Chloroform is observed between 0.62 and 0.75 GPa [13], whilst δ-bromoform [13] crystallises at much lower pressure of 0.2 GPa and iodoform transforms at 0.85 GPa from a disordered $P6_3/m$ structure where the mirror symmetry is present through the iodine atoms to the ordered $P6_3$ structure where all the molecules are aligned [84] (Fig. 8). The transformation from the disordered model to the ordered model can also be observed on cooling (bromoform has an abrupt change at 270 K, whilst iodoform shows a gradual change) [12]. The packing in the $P6_3$ phase allows the recognisable edge to end interaction between the halogen atoms that is the result of the atomic charge distribution of the atoms.

Substituted benzenes have also been the focus of studies into halogenated van der Waals materials. Ridout and Probert investigated the three isomers of monofluorotoluenes using the low-melting techniques described earlier, but in this study, the pressure was cycled rather than crystal growth through heat annealing [85]. This type of cycling requires that the sample is close to the liquidus line and that the solidification is directly into a new phase. The high-pressure forms of 3- and 4-fluorotoluene follow a similar trend to the alcohols where the intermolecular interactions are sacrificed to maximise the packing efficiency of the aromatic groups. The authors observed some interesting behaviour of 2-fluorotoluene that mimics the behaviour that we observed in 2-chlorophenol/2-methylphenol system [21]. They did not observe a new form on compression but only through a failed co-crystallisation attempt with 3-fluorotoluene. The new high-pressure polymorph was observed at slightly higher pressures than the original phase was observed. The authors note that the mixture permitted the growth of the thermodynamic product due to slow nucleation. In addition to this, the observation of the new phase may be attributed to the fact that 3-fluorotoluene was acting as the 'solvent', due to its lower melting point, allowing a higher pressure to be achieved. Through the mixed PTM method, the authors were able to cross the phase diagram boundary between the LT

and HP phase and isolate the new form. In each of the phases, the CH. . .F interaction is favoured.

Anioła et al. have probed chlorinated and bromonated meta-disubstituted benzenes by mixing the material with methanol to help to achieve higher pressures than would be able using the pure forms similar to our previous example (mixed PTM method) [86]. m-Dichloro- and dibromobenzene were investigated individually as well as a solid solution. The pure phases observed at high pressure were the same as those crystallised under cooling. An interesting observation from a methodological point of view was that in an attempt to co-crystallise the two solids together, solid solutions are formed from methanol solutions but only at pressures greater than 1.02 GPa; below this the pure m-dichlorobenzene phase is favoured. If, however, a mixture of the two pure materials is used, without methanol, then the solid solution can be isolated. In this case, the difference in the solubilities of the pure materials in methanol is an important property at lower pressures, i.e. the m-dichlorobenzene precipitates before the dibromobenzene, but at higher pressures the difference in this physical property becomes negligible; hence, the solid solution can be crystallised.

The latter study of m-dichlorobenzene, together with the 2-chlorophenol [21], 2-fluorotoluene [85] and the acrylic acids [31, 33, 34], demonstrates that even if there are no apparent phase transitions at high pressure, there may be potential for the isolation of new forms through alternate methodologies. It is worthwhile to explore mixing the compound of interest with another solvent or PTM to investigate whether the system can be pushed to higher pressures and potentially over phase boundaries to new unidentified phases.

3.2.2 Compression of Solids

The next few examples explore the effect of pressure on heteromeric intermolecular interactions whether they are in a single pure compound or whether they are in a multicomponent form (salt or co-crystal).

4-Iodobenzonitrile is a simple disubstituted benzene that possesses an intermolecular interaction between a nitrile group and the iodine [87]. The solid is constructed through chains of molecules, exploiting this interaction, that lie antiparallel with their neighbours. The compression of this phase demonstrates the importance of the non-directional π. . .π and weak H. . .N interactions rather than the visually recognisable N. . .I halogen bond. Giordano et al. observed that the π. . .π interaction becomes very repulsive on compression until a phase transition occurs at 5.5 GPa from monoclinic *I*2/*a* to *P*-1. The lower symmetry enables the separation of the π. . .π interaction into two symmetry independent interactions so that one of these interactions increases in length, stabilising it with respect to the lower pressure phase (21 $kJmol^{-1}$ cf. 28 $kJmol^{-1}$), whilst the other continues to be compressed (Fig. 9). The authors identified that it is the change in these ‘weaker’ interactions that promotes the phase transition at pressure. So even in these halogenated compounds, the packing efficiency is key, sacrificing the length of the more recognisable N. . .I interaction for a closer packing of the main benzene backbone.

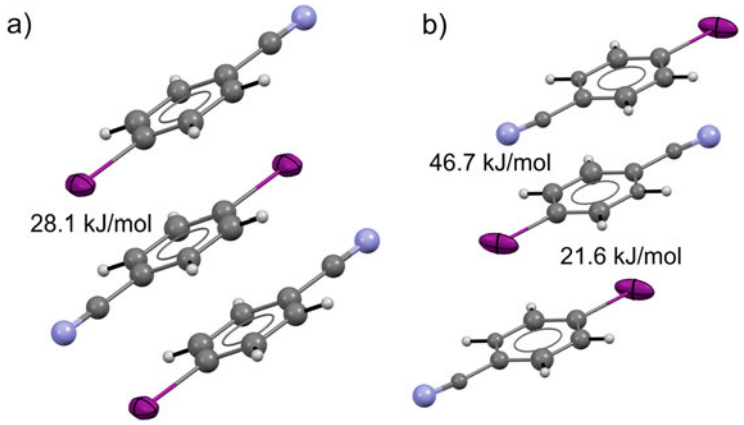


Fig. 9 (**a**) π... π interactions between molecules in Form I of 4-iodobenzonitrile indicating the energies calculated using PIXEL. The energy of interaction between the central molecule and those above and below is the same. (**b**) π... π interactions in Form II of 4-iodobenzonitrile indicating the disparity of interaction over the phase transition. One of the π-interactions becomes more destabilising than the other

Cinnamic acids have been a great source of investigation of [2 + 2] cycloadditions [88, 89]. Whilst these papers investigated the role of irradiation and temperature on the reactivity of these compounds, the group of Turowska-Tyrk utilised pressure in addition to UV irradiation to investigate a set of halogenated cinnamic acids (2,5-difluorocinnamic acid, 3,5-difluorocinnamic acid [90] and 2,6-difluorocinnamic acid [91]). In each of these structures, the CH. . .F interaction played a role in aligning the functional groups to enable the [2 + 2] cycloaddition to occur. They observed that by the application of pressure, the rate of the reaction increased over those experiments conducted at ambient pressure but in general the reaction only went to a maximum of 50% completeness. One of the interesting observations was that in 2,5-difluorocinnamic acid the reaction rate maximised after 0.3 GPa; reactions at 0.9 GPa proceeded at a similar rate. The reactions of 2,6-difluorocinnamic acid only went to 35% completeness at 2.1 GPa after 270 s of irradiation. At higher pressures the molecules are not as mobile (unless undergoing a phase transition); hence, the reaction is suppressed. This would help to support this hypothesis of Kaupp who believes the ability of a structure to move is key to polymerisation rather than a set distance [92]. This is in contrast to the study of 2,4,6-tricyano-1,3,5-triazine where shorter intermolecular contacts were identified as the driver to reaction as opposed to the crystal structure dynamics [93].

In an excellent piece of crystal engineering, the work of Goroff and co-workers exploited halogen bonding to form a multicomponent crystal form of diiodobutadiyne with a selection of oxalamides to explore polymerisation using high pressure [94, 95]. The halogen bonding was used as a tool to direct the molecules into adopting a geometry where the dyene moieties were in close proximity which may not have been possible using a single component system. Their

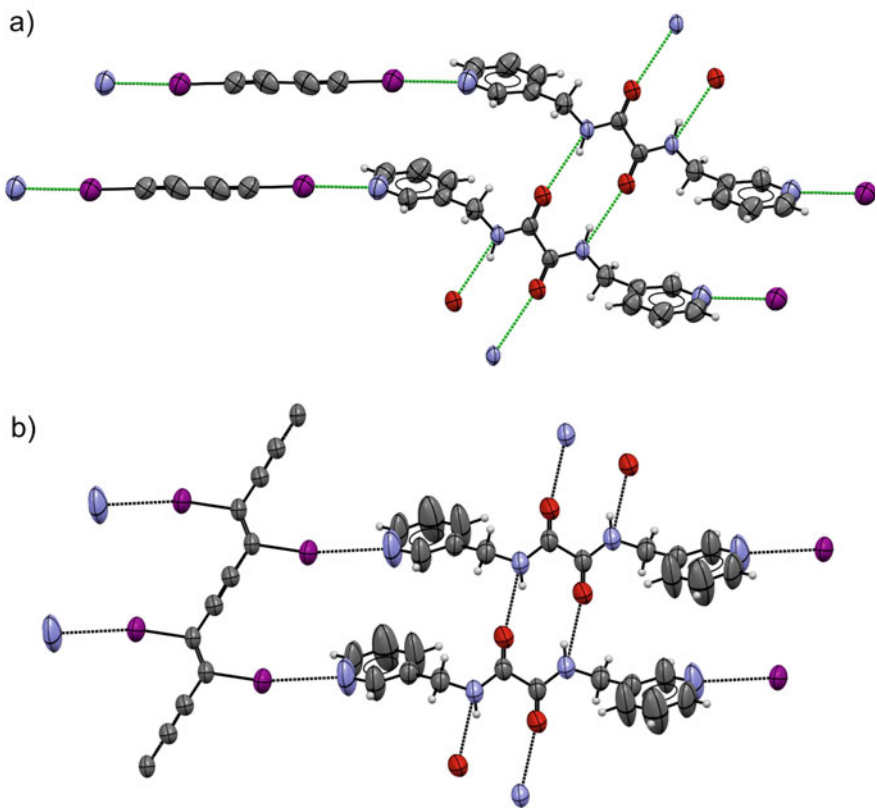


Fig. 10 (**a**) Bis(pyridine)oxalamide: diiodobutadiyne co-crystal at ambient pressure indicating the alignment of the polymerisable groups through formation of a co-crystal via hydrogen and halogen bonding; (**b**) the polymerised product after compression to 3.5 GPa. The pyridine moieties indicating a tendency to a more disordered position

previous work had shown that using bis(nitrile)oxalamide as the co-former, the polymerisation of diiodobutadiyne is enabled at ambient temperature and pressure; however, it was hindered using pyridine-based oxalamides due to steric effects. In two papers, the authors were able to demonstrate that by application of pressure, the polymerisation could be achieved using the sterically hindered pyridine molecules. The reaction could be observed through a colour change in the solid as pressure was applied to the system which was then correlated to the reaction occurring in the co-crystal monitored using X-ray diffraction (Fig. 10). The authors note that the polymerisation begins almost immediately as pressure is applied to the system. Unfortunately, there is only the polymerised structure available and no indication whether subtle phase transitions occur during the compression as we observed on compression of acetylenedicarboxylic acid [25]. Acetylenedicarboxylic acid undergoes two phase transitions prior to polymerisation at 5.2 GPa, or more

accurately oligomerisation that may have helped to initiate the reaction following the hypothesis of Kaupp.

The theme of reactivity in halogenated compounds has recently been continued through the study of a charge-transfer complex, tetraethylammonium diiodine triiodide salt [96]. The salt was selected from the Cambridge Structural Database [97] due to the favourable interaction of the iodine groups in the crystal structure. The authors were able to show that by the use of pressure, the iodine groups, I_2 and I_3^-, interact to a greater extent so that the salt becomes increasingly conductive as pressure is applied with a distinct increase in conductivity between 9.4 and 11.1 GPa. From the crystal structures, the authors were able to extract intermolecular information, thereby elucidating the nature of the bonding between the iodine groups that they identified as a mixture of ionic and covalent bonds.

3.3 Amino Acids

No chapter on high-pressure crystallography would be complete without mention of amino acid chemistry. There are a number of excellent reviews on amino acid chemistry at high pressure provided by Boldyreva [8], Moggach et al. [2] and Freire [98]; hence, we are going to highlight more recent studies that include those that have pushed the boundaries of what can be explored using small molecule systems. Table 1 provides a non-exhaustive list of amino acids that have been investigated with pressure using diffraction along with their references.

Studies of alanine were the first to have pushed the pressure boundary for amino acid structural chemistry. The two studies by Tumanov et al. and Funnell et al. used complementary techniques to explore this system. Funnell et al. [104] used neutron powder diffraction to follow the structural changes in L-alanine before compressing it into an amorphous material at 15.46 GPa. The authors were able to follow the compression of the crystal structure to 13.6 GPa and demonstrate that L-alanine remains in the same phase over this pressure range which was in agreement with Tumanov et al. [103] who reached 12.3 GPa using X-ray powder diffraction. The reduction in voids in crystal structures has been noted before as a reason for

Table 1 Single-crystal X-ray studies of various amino acids

Amino acid	References
Glycine	α- to ε-glycine [99, 100] (δ-glycine [101]), ζ-glycine [102]
Alanine	L-Alanine: Form I [103–105], D,L-alanine: Form I [106]
Cysteine	L-Cysteine: Forms I–IV [107], D,L-cysteine [108]
Serine	L-Serine: Form I, II [109, 110], III [111] and IV [112], monohydrate [113], D, L-serine: Form I [114]
L-Threonine	L-Threonine Forms I–III [115]
L-Glutamine	L-Glutamine: Form I [116]

compression or phase transitions [107, 109, 117]. In this case, L-alanine was compressed to an extreme where only 11.11 Å of the total cell volume was attributable to voids at 13.6 GPa and where the void volume has reached a minimum before amorphisation is induced.

The latest study of an amino acid, L-threonine, has achieved the greatest pressure to date for these materials [115]. The study of Giordano et al. demonstrated that L-threonine can be compressed to 22 GPa and still achieve molecular level detail of the changes that occurred in the system. They observed that L-threonine undergoes three phase transitions at ~3.2, ~9 and ~17.5 GPa of which the initial two transitions are isosymmetric. The reasons for the transformations are all different with the first transition due to a molecular rotation of the carboxylate group; the second indicates a change in the compression behaviour of the phase, whilst the last is driven by the reduction in the molecular volume through conformational change in the hydroxyl group. Staggeringly, despite the pressures achieved, the authors noted that the hydrogen bonding present remained within the bounds of the normal distribution that is observed under ambient conditions using data from the Cambridge Structural Database. Reference to the CSD mean values has been used previously to rationalise polymorphic transitions in molecular systems [109].

3.4 Pharmaceutically Relevant Materials

In this next section, we will explore the pressure dependence of drug compounds and compounds related to pharmaceutical processing. The increase in hydrogen bonding groups and flexibility make them particularly susceptible to polymorphism which is one of the greatest challenges for the pharmaceutical industry. To this end polymorph screening is a regular process during the drug discovery pipeline. During a polymorph screen, different solvents and crystallisation conditions are explored under ambient pressure conditions, but the variation of pressure is not used. As we have observed, the variation of pressure can have a marked effect on the crystallisation outcomes of fairly simple molecules; hence, it is intuitive that more complex molecules with a wide variety of hydrogen bonding groups and flexibility will form new polymorphs at elevated pressure. Many groups have investigated pharmaceutical materials, and here we discuss a selection of them to highlight some of the key findings and/or experimental procedures that have helped to characterise the new forms (Fig. 11).

3.4.1 Chlorothiazide (I)

Chlorothiazide is a diuretic and can help to reduce swelling caused by cases of kidney or liver diseases. The interest in chlorothiazide stems from it being a pharmaceutical product but also that it is a relatively rigid molecule. The rigidity of the molecule was important at the time of the study due to the difficulty of solving

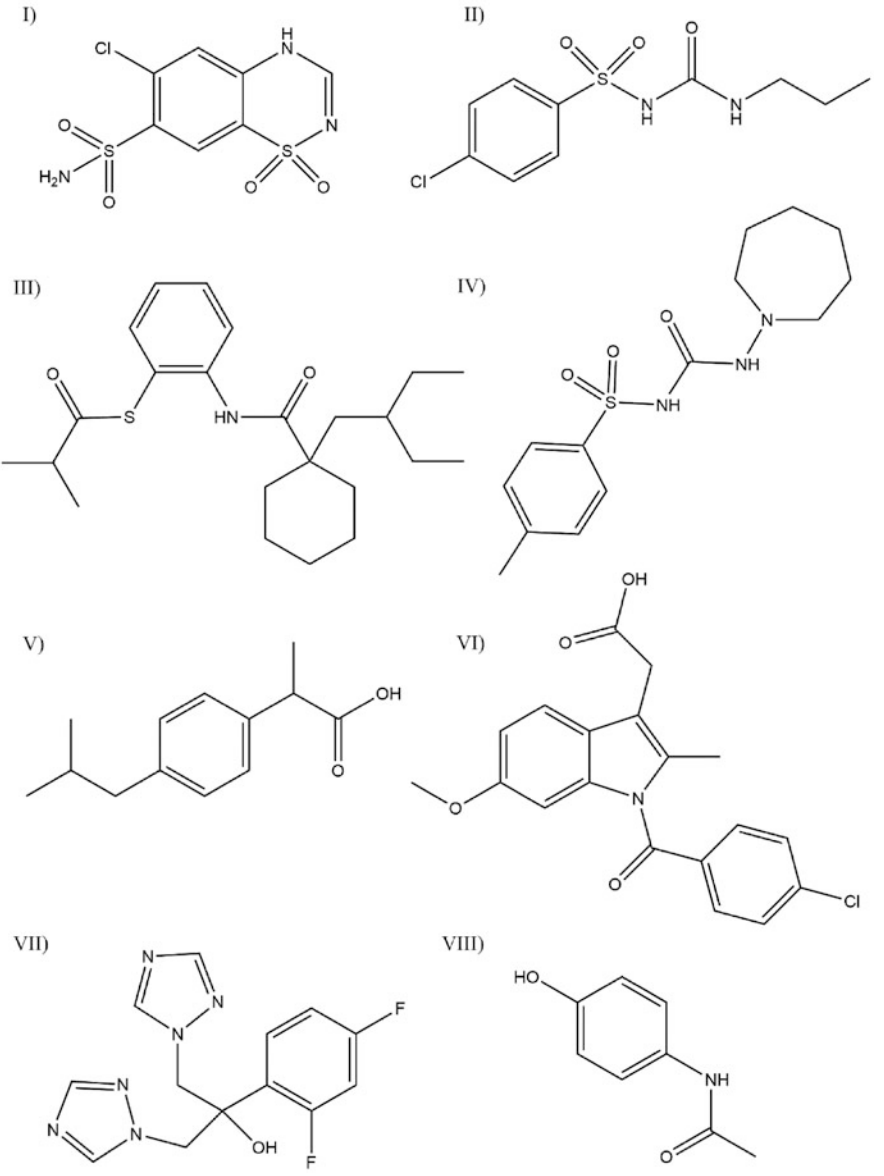


Fig. 11 Chemical diagrams for some pharmaceutical materials investigated using pressure. (I) chlorothiazide; (II) chlorpropamide; (III) dalcetrapib; (IV) tolazamide; (V) ibuprofen; (VI) indomethacin; (VII) fluconazole; (VIII) paracetamol

new crystal structures with the incomplete data that come with using the DAC. This was a particular challenge due to the low symmetry of the known ambient pressure form (P1) [1]. Real-space structure solution packages such as DASH and FOX were

key to the solution of a number of high-pressure phases including RDX [118] and paracetamol methanolate [41]. In this study the authors used multiple crystals to increase the proportion of the reciprocal lattice that could be observed to 69% compared with a typical dataset completeness for a triclinic system being ~30%. Chlorothiazide demonstrates one of the key features of high-pressure studies whereby the torsional angles of the molecular systems are the first changes to molecular geometry that take place. At 4 GPa, chlorothiazide undergoes an isostructural phase transformation where the thiazide ring is distorted due to the proximity of the intermolecular interactions between the SO_2 moiety and the ring system of a neighbouring molecule; the sulfonyl group also shows a large conformational change due to the shear of the molecular planes in the crystal. The phase transition is reversible so it is not particularly useful in terms of the polymorph discovery, but the study demonstrates methodologies that have been used to counteract the issues around data completeness and structural solutions of systems at high pressure albeit the latter has somewhat been overcome by newer solution packages, e.g. ShelXT [119].

3.4.2 Chlorpropamide (II)

Chlorpropamide, (4-chloro-N-(propylaminocarbonyl)benzenesulfonamide), is one of the best examples of how high pressure may affect the nucleation and growth of high-pressure phases. It is a complex molecule used as an anti-diabetic and possesses five different polymorphs (α-ε) that can be obtained under a range of crystallisation conditions under ambient conditions. The α-form is the most thermodynamically stable polymorph but does not possess the highest density. On cooling the β and ε-forms, the structures undergo transitions to structurally similar low-temperature forms, β^{II} & β^{III} and ε', respectively [120, 121]. The α-δ-forms have been shown to transform to the ε-form near the melting point [122]. The extensive polymorphic behaviour at ambient pressure suggests that the application of pressure would enable the transformation to further high-pressure polymorphs. In a range of papers exploring the effects of pressure, the group of Boldyreva have observed a number of new phases and some interesting effects related to the PTM used.

During their studies Boldyreva et al. employed a number of different high-pressure techniques from simple compression of a dry powder through to the alteration of the pressure-transmitting media used for the study. We have already touched on many of these concepts in this chapter though this example demonstrates the choice of experiment available to investigate molecular forms. In this example, Boldyreva et al. showed that pure compression of a system can induce a polymorphic transition in the case of both the α-, β- and δ-forms. In one of the first studies, they investigated the α-γ transition that had previously been reported to occur on compression during the tabletting procedure [123]. The α-form crystallises in orthorhombic $P2_12_12_1$ with $Z' = 1$. Initial X-ray powder diffraction experiments showed little change to the dry powder on compression except a general broadening of the pattern through the non-hydrostatic compression. There were small inconsistencies

in the unit cell parameters above 3.7 GPa that indicate the transition to the new phase as well as small shoulders on some peaks that may support this conclusion. A follow-up study used single-crystal diffraction to investigate the compression so that structural changes could be followed as a function of pressure as well as to provide conclusive evidence of any phase transitions [124]. For this experiment ethanol was used as the PTM to ensure hydrostatic compression. The addition of a PTM is a necessary and may alter the behaviour on compression. However, in this case, the transition could be observed into a new monoclinic ($P2_1$) structure with $Z' = 2$ at 2.8 GPa (designated α'); the crystal of the high-pressure phase is twinned due to the reduction in the symmetry. The authors also highlight the difference in the transition pressure between the powder diffraction experiment and the single-crystal experiment and attribute this to the presence of nuclei in the crystal. For a transition in a single crystal, there only needs to be one nucleation point for the entire sample to transform. For a powder, every particle requires a site of nucleation that may delay the onset of the transition. From a structural perspective, chlorpropamide shows that the packing of the molecules takes precedence over the hydrogen bonding; cf. simple alcohols. Of the three hydrogen bonds present, two increase in length to accommodate the packing of the molecules where there is a change in the torsional angles around the phenyl and alkyl groups.

Boldyreva et al. continued their exploration of chlorpropamide in an interesting competitive experiment where they investigated the role of nucleation and seeding (Fig. 12) [125]. For this experiment, they chose a PTM (1:1 pentane/isopentane) in which chlorpropamide was visibly 'insoluble'. In fact, they observed a difference in the solubility of each polymorph where the α- and δ-forms were visibly unchanged but the metastable β-form showed rounding of the edges of the crystal. This is a critical observation as the slight solubility of the crystal impacts on the behaviour of β-form at high pressure which was demonstrated in a later study [126]. The solubilisation of the β-form in the PTM allowed the transformation to the γ-form alone without any seeds present or to γ- and δ-form if seeds of α- and δ-forms were present. The transformation to γ-form was rationalised due to the similarity in the packing between the β- and γ-forms, hence providing a low barrier to interconversion. The authors conclude that in solution molecular clusters are retained which is aided by the confined geometry of the DAC sample chamber. Transition to the α-form is inhibited due to the substantial rearrangement required. The increase in the δ-form was attributed to the seed crystal providing a template that enables the facilitation of the transition despite the large molecular rearrangement required. It is also the densest phase; hence, the pV term of the Gibbs free energy equation will provide a thermodynamic driving force.

In their final study of chlorpropamide (at time of writing), the Boldyreva group explored the properties of the PTM further using β-chlorpropamide. β-Chlorpropamide showed slight solubility in the 1:1 pentane/isopentane mixture; hence, the authors explored the use of PTMs in which there was going to be no solubility, e.g. helium, neon and paraffin oil [126]. They observed very different behaviour depending on the use of each of these media. In neon, they observed transition of the β-form to the α-form over the course of 2 days at 0.6 GPa. The

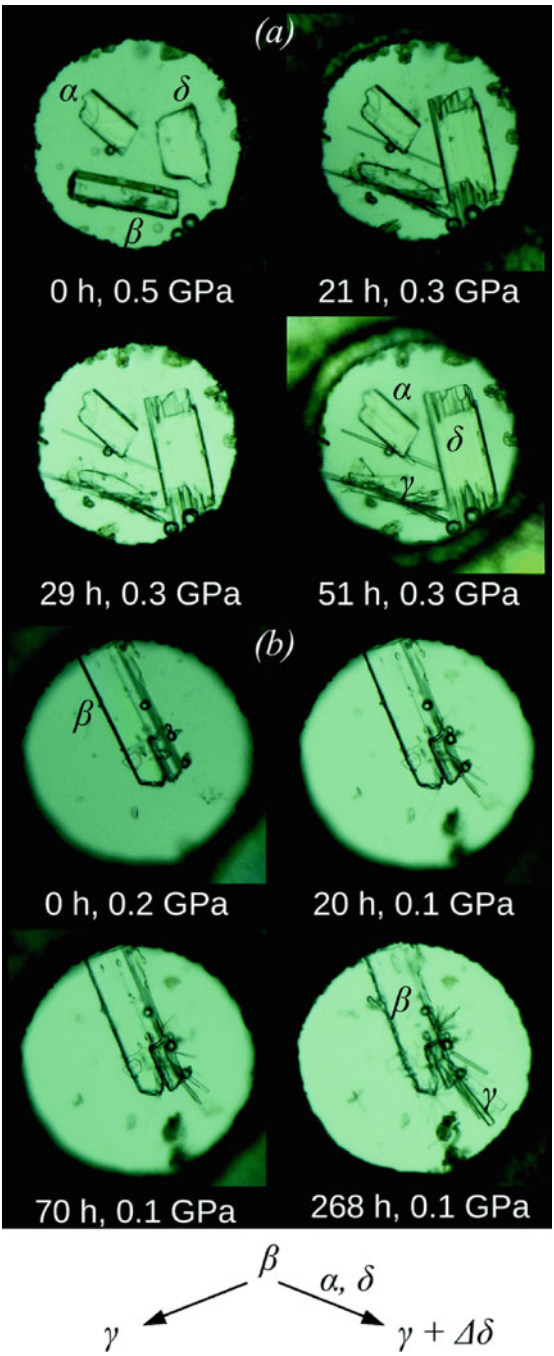


Fig. 12 Seeding experiment conducted by Boldyreva and co-workers indicating the nucleation of different phases depending on the initial phases present. Figure taken from CrystEngComm, 2016, 18:5423–5,428, DOI: (https://doi.org/10.1039/C6CE00711B) with permission from Royal Society of Chemistry [125]

substantial molecular rearrangement caused a reconstructive phase transformation into multiple smaller crystals. Further compression of this phase showed a transition that they speculated to be the α'-polymorph that they observed previously, but the diffraction was not of sufficient quality for unequivocal assignment. In helium and paraffin oil, the β-form transforms to a structurally similar phase β^{I}_{HP} at 0.3 and 0.1 GPa, respectively, but there were differences in the behaviour of the crystals. In the helium medium, the single crystal was retained; however, in the paraffin oil, the crystal fractured into several new domains. Further compression in helium to 0.7 GPa initiated a change to β^{II}_{HP} where the crystal fractured over the transition. There are similarities to the low-temperature phases, but they could not be fit to the data. In paraffin oil, there is a transition to a triclinic phase that is designated β^{III}_{HP} with further fracture of crystal between 0.1 and 0.3 GPa. The very distinct behaviour in each of these media was explained by the interaction of the media with the crystal itself. In the case of helium, it is known to enter the structures of 4-hydroxycyanobenzene [49], silicates and even diamond itself (the cause of many broken diamonds at very high pressure). The phase transitions are facilitated by a combination of internal and external pressure. Paraffin oil interacts with the surface promoting the change in the polymorph. The fact it is a collection of alkyl molecules and that pentane/isopentane mixture interacts with the surface may allow that hypothesis to hold true. For neon, the size of the atom will exclude the penetration into the crystal structure; hence, thermodynamics may be playing a more significant role in this transition. Unfortunately, the authors were not able to perform the loading again to investigate the kinetics of the transition further.

3.4.3 Dalcetrapib (III)

This example typifies the strategy that has been explored to help to inform the discovery of new polymorphs of pharmaceutical materials. The study of Neumann et al. [6] explores the combination of crystal structure prediction together with the high-pressure techniques to explore the energy landscape of dalcetrapib. Typically, the energy landscapes calculated ab-initio reveal hundreds of crystal structures that are energetically reasonable. However, depending on the compound, only a handful of these may have been characterised through experimental procedures. Density or packing coefficient is usually plotted along with the energy of the structures; hence, it is a very clear method to identify potentially new unobserved forms that may be more thermodynamically stable than the known forms of a drug. This is a red flag for pharmaceutical companies, but if one has exhausted ambient pressure polymorph screens, there is a possibility that pressure can be used to access new forms. This study used recrystallisation at high pressure to isolate a new polymorph of dalcetrapib from tetrahydrofuran. At very low pressures, a new polymorph was identified and compared well with one of the predicted forms albeit it was not the most stable global minimum structure. The authors attempted to recover the sample, but it was thermodynamically less stable than the known form and converted over the course of a few hours.

3.4.4 Tolazamide (IV)

At ambient pressure tolazamide possesses three polymorphs, two of which crystallise concomitantly (Forms I and II). Form I is an ordered structure and is suggested to be the thermodynamically stable form, whereas Form II displays disorder around the azepane ring. The energies between the two phases are similar, but Form II converts to Form I in solution. Form III was isolated once from a hot dioxane solution but never obtained again and hence may require very specific experimental conditions to crystallise [127]. The study of Fedorov et al. investigated the pressure effects on both Forms I and II and explored the use of different PTM building on the work on chlorpropamide [128]. This is another case where there is a kinetic barrier to the transition due to the molecular rearrangement required. Using pentane/isopentane as the PTM, both of the polymorphs can be characterised to 6 GPa without any indication of phase transitions. The density of Form I is higher than Form II indicating the stability of this phase over the entire pressure range and beyond 25 GPa (through computational approaches). The bulk moduli of the two phases are similar with Form II being slightly more compressible (Form I 6.4(3) GPa and Form 5.8(2) GPa). This is reflected in the hydrogen bonding where they compress more rapidly in Form II than in Form I. In addition, the disorder in Form II is 'frozen out' at 3.3 GPa. This also aligns well with many other studies of molecular materials at pressure that show a reduction in void volume [109]. In this system it can be surmised that due to the disorder present, the structure possesses a greater volume of void space that can be compressed; hence, there is a lack of phase transformation; we have noted this in the compression of 3,4-methylenedioxymethamphetamine hydrochloride [129].

The exploration of tolazamide in methanol indicated a change in Form II to Form I at 0.1 GPa which is observed at ambient pressure. Whilst the change in the phase cannot be attributed to the pressure, the conversion was not fully complete when conducted in the DAC. The authors do not specify the length of time that the sample was left at 0.1 GPa, but they do conclude that the lack of solubility in methanol prevents the complete conversion. We have also noticed this phenomenon in *p*-aminobenzoic acid where we believe that *p*ABA is saturated in the medium and the viscosity of the solution prevents the conversion to the new phase.

3.4.5 Other Pharmaceuticals

Ibuprofen (V) is a well-known chiral non-steroidal anti-inflammatory. Its racemic form is known to exist in two polymorphic forms. Form I is the most stable polymorph with the second polymorph being observed only after quench cooling and annealing at 143 K for an hour and 258 K for 15 h [130]. Pure (s)-(+)-ibuprofen is only observed in one polymorph [131]. Ostrowska et al. investigated the racemic compound using both recrystallisation and compression methodologies to a pressure of 4 GPa with a view that it may be chirally resolved using pressure [132]. Despite the use of many solvents (chiral and achiral) and range of pressures and temperatures (0.1–1.15 GPa; up to 553 K), the known Form I was always produced. Even attempted co-crystallisations with mandelic acid and tartaric acid failed. The idea

of chiral resolution with pressure has been investigated by Rietveld and co-workers who investigated mandelic acid. They observed that the chirally pure forms of mandelic acid are stable above 0.64 GPa and 460 K based on calorimetry [133]. This has since been disputed by two articles that explore the mandelic acid system in its chirally pure and racemic forms. The thermal event that was observed by Rietveld and co-workers was identified as the transition between the low pressure and high pressure phase of rac-mandelic acid [134, 135].

Indomethacin (VI) is a well-studied pharmaceutical due to the polymorphism that it exhibits. It has three solid forms, a metastable α-polymorph, stable γ-polymorph as well as an amorphous form. Okumura et al. explored the crystallisation behaviour of indomethacin to 0.4 GPa to observe whether pressure would alter the phase behaviour or induce amorphisation. They discovered that as a dry powder both polymorphic forms of indomethacin were reduced at the expense of the amorphous form (ca. 50% content to ca. 20%) at 0.4 GPa. As a slurry, the behaviour was slightly different as recrystallisation from the solvent was possible to induce a change from the γ-form to the α-polymorph at 0.4 GPa. There was still some transformation to the amorphous form. The authors reasoned that the change in behaviour was due to the increased density of the alpha form as well the increase in solubility of the amorphous form, but they noted the need for solubility measurements at pressure to confirm this. The preparatory routes for the X-ray powder diffraction in this study have obscured the output phases due to the grinding in a mortar and pestle prior to data collection. This may have helped to initiate the amorphisation. Nevertheless, this study contributes to the overwhelming observation that solvation can help to overcome the kinetic barriers to interconversion at high pressure. We have followed on from this study by investigating the co-crystal of indomethacin and saccharin at pressure [136]. The co-crystal was stable over the entire pressure study to 6 GPa and was even reproducibly formed via recrystallisation at pressure. This compression study allowed us to investigate the effect of pressure on different types of hydrogen bond, but it was the packing of the molecules in a host/guest manner that had an effect over its compression. By its placement in a guest environment, the saccharin molecules could hydrogen bond together at a distance close to a gas-phase minimum due to the packing of the large bulky host, indomethacin.

Fluconazole (VII) is an antifungal agent that shows extensive polymorphism albeit the polymorphs have not been fully characterised [137, 138]. There are seven forms of fluconazole that are speculated to exist albeit that crystal structure of Form I has not yet been determined and Form III is only speculated to exist due to inconclusive spectroscopic data. Gorkovenko et al. investigated the pressure dependence of Form I using energy-dispersive X-ray diffraction and observed the isolation of a new triclinic phase (Form VIII) at 0.8 GPa with the possibility of a second transformation at 3.2 GPa. From the change in the isotropy of compression of the unit cell parameters, the authors speculated that the transition to the new phase would require a substantial rearrangement of the molecules. Fluconazole only possesses one hydroxyl group that hydrogen bonds to the triazole group in the known crystal forms or to water molecules in the hydrate. This molecular structure gives rise to bulk moduli closer to glycolide [65], caprolactam [139] or aniline [22] than other drug substances.

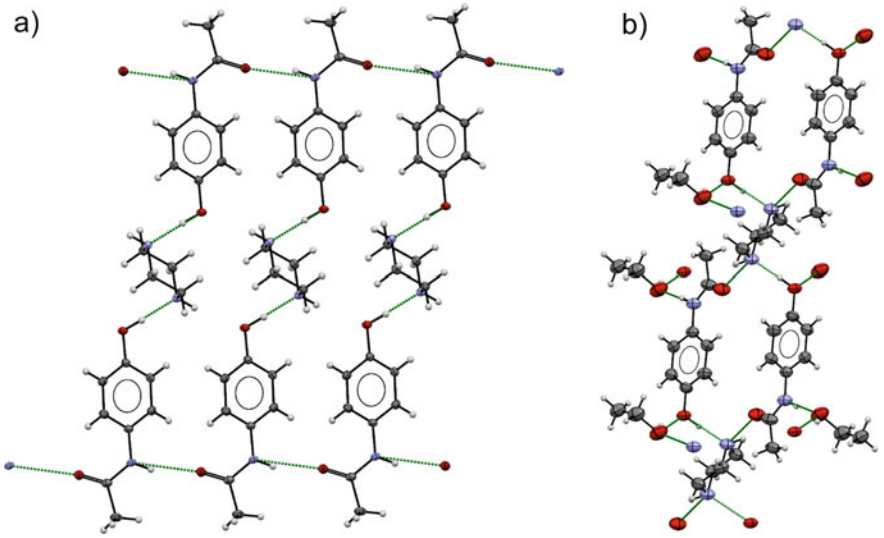


Fig. 13 (**a**) The ambient pressure co-crystal of paracetamol/piperazine; (**b**) the crystal structure of the paracetamol/piperazine/ethanol co-crystal solvate grown at 0.57 GPa showing the change in the hydrogen bonding between the molecules. This phase has an unusually compressed OH. . .N hydrogen bond compared to ambient pressure observations in the CSD

Co-crystals of paracetamol (VIII) have also been investigated at pressure. The paracetamol/piperazine system was used as an exemplar to demonstrate the potential to isolate new forms of co-crystals at high pressure. In this study we were able to use the recrystallisation technique to isolate a new ethanol solvate of paracetamol/piperazine at 0.57 GPa [140]. The co-crystal solvate hydrogen bonds in a very different manner. The hydroxyl to amine interaction is still present, but the piperazine now hydrogen bonds to the amide group of a neighbouring paracetamol. The OH. . .N interaction is compressed by 6.5% over the mean interaction in the CSD which is remarkable given the small pressure change. The ethanol hydrogen bonds to the paracetamol hydroxyl group (Fig. 13). Like many of the solvated systems, the solvate was non-recoverable and converted to the known co-crystal which is not surprising given the strong perturbation of the hydrogen bond length. Whilst the discovery of this solvated co-crystal may seem routine, it does highlight the potential to access new phases at pressure that can be desolvated on decompression. In the future, this may provide a route to access new polymorphic forms that are thermodynamically stable at ambient pressure.

3.4.6 Illicit Materials

We began a series of studies investigating the role of pressure on illicit substances. This was initiated with the wave of new ‘legal highs’ that were beginning to flood the

marketplace for club goers. The rationale for investigating materials using pressure to this point was to find new polymorphs of pharmaceuticals, but the area of illicit materials had been overlooked despite the unregulated nature of production of these materials; the potential for the discovery of new forms was and is still very high.

Mephedrone was one of the 'legal highs' of choice hence we investigated the hydrogen sulphate salt of this material [141]. We observed that mephedrone hydrogen sulphate was quite polymorphic on compression and that the phase transitions to the two new high-pressure polymorphs were single-crystal to single-crystal which enabled ease of characterisation. The first phase transformation was between 0.5 and 0.88 GPa and was isomorphous with a compression of the unit cell volume by 10%. This was enabled by the change in the orientation of the sulphate chains, whilst there was a small shift in the organic counterion. The second phase transition occurs between 3.56 and 4.8 GPa with a reduction in the symmetry to *P*-1 and increase in Z' to 4 which made the refinement of the model challenging. This phase transition is characterised by a marked reorientation in the sulphate chains but also in the torsional angles in the mephedrone molecule itself. In many of the studies at pressure we have observed that hydrogen bonding lengths are limited to the ambient pressure values in the CSD but this case we were able to demonstrate that this in also the case for phenyl-phenyl interactions. Prior to the phase transition, the phenyl groups of the mephedrone molecule are compressed to the lower limits observed in the database before the transition. After the transition two out of the four interactions move substantially away from the limit.

We followed with a study of 3,4-methylenedioxymethamphetamine hydrochloride or MDMA [129]. Disappointingly, there were no phase transitions observed in this system; however, one of the key findings was that due to the orientation of the molecules with respect to each other, there was a correlation with respect to the principal axis of strain where the compression of one axis limited the compression along another like a wine rack effect which had been observed previously in relation to metal-organic framework materials [142]. Furthermore, the orientation of the hydrogen bonds allowed for the structure to be compressed easily in that direction as they acted like a spring which may have attributed to the system not showing any polymorphism with pressure. We have recently noted this spring-like effect in the compression of two high-pressure polymorphs of ε-caprolactam [139]. The two forms have similar hydrogen-bonded chains; however, the interconnectivity of neighbouring chains is different. We noted a large compressibility along the chains where the neighbours were not intercalated, but this was reduced by half in the polymorph where there was intercalation.

Overall, we see a wide range of behaviours in organic molecular compounds under high-pressure conditions. The overriding message from this is that the interaction of the PTM can have a huge impact on the ability to observe new phases of molecular forms. Hence the planning and preparation for high-pressure studies is crucial for a successful outcome. The high degree of flexibility can create issues for the molecular transformations, and the kinetics of such can be very slow or suppressed completely. Whilst we have only provided a few examples of molecular materials at high pressure, these have been largely focussed on single-crystal

diffraction studies where the various groups have been able to rationalise changes to structure through following the molecular movement as a function of pressure. There are many more studies using X-ray powder diffraction, Raman and IR that will provide readers with a significant body of research to digest. That being said, there are many more materials out there to be squeezed and probed with X-ray diffraction methods, and we hope that we have provided, at least, a starting point from which to begin your studies.

4 The Effect of Pressure on Metal-Containing Complexes and Framework Materials

4.1 Introduction

The effect of pressure on metal complexes, frameworks, coordination polymers and metal-organic framework (MOF) materials has been extensively reviewed, and we would particularly like to point those of Moggach and Parsons [143], Tidey et al. [144] and Gütlich et al. [145]. These reviews comprehensively cover the effect of pressure on metal complexes and spin-crossover materials, whilst the review by McKellar and Moggach [50] discusses the effect of pressure on MOFs. We will not be discussing in detail effects such as negative linear compressibility (NLC behaviour), where an expansion along a particular crystallographic direction is observed on increasing pressure [146], despite this effect being observed in a number of metal-containing compounds recently. If the reader is interested in NLC behaviour, we encourage them to read much more thorough and focussed reviews, such as the excellent review by Cairns and Goodwin [147].

In this section, we will provide an overview of the science of metal-containing compounds under high-pressure conditions conducted since the last review of the area in 2015. As with the previous section on organic materials, metal-organic materials are a broad and keenly studied area, and we have tried to highlight a few topics that may be of interest to the reader.

One of the stark differences between the compressibility of organic compounds and metal-containing materials in the solid state is the compressibility of the metal-ligand (M-L) bonds which are far more compressible than covalent bonds. For example, Cu-ligand bond distances can vary from 1.9 Å to 3.1 Å and still be considered a bonding interaction. As a consequence, far greater changes in intramolecular bonds can be observed. Surprisingly, not only do we see the occurrence of bond formation in metal complexes but also the ability to break bonds on increasing pressure. This ability to tune the formation and breakage of bonds is limited to specific sets of organic materials that possess the correct functional groups to facilitate the transition and hence is not so widely applicable (Sect. 3.2.2). Just as in the organic solid state though, conformational changes occur, whilst changes in intermolecular interactions in compounds that contain a metal often give rise to

spectroscopic changes. The emergence of high-pressure diffraction studies, conducted in parallel with other physical property measurements, such as luminescence, fluorescence, and magnetometry, has really opened up the possibility of deriving structure/property relationships in a plethora of functional materials. Here, we will give examples of recent studies that show changes in intramolecular conformation, M-M, M-L, and intermolecular interaction distances, which can lead to both the formation and breakage of bonds. Finally, we will give examples of how these changes have been used effectively to develop structure/property relationships in molecular magnetic materials, spin-crossover complexes and, finally, metal-organic frameworks and coordination polymers.

4.2 *Intramolecular Conformational Changes and Compressibility of M-M and M-L Bonds*

The investigation of the coordination environment has been a source of rich discovery from purely structural context as well as from a structure/property perspective. The decrease in length of metal-metal bonds has been observed in a handful of metal complexes. For example, in the linear chain compound $Co_2(dpa)_4Br_2$ (dpa = 2,2′-dypyridylamide anion), a reduction in unit cell volume of 30% is observed at 13.6 GPa. Most of the compression is 'taken up' by the compressibility of much softer intermolecular interactions that cause the redistribution of DCM (dichloromethane) solvent molecules during the compression. However, a portion of the compression is due to a decrease in length of the Co-Co and Co-Br bonds with each showing a reduction of 4% and 12%, respectively, at 11.8(2) GPa [148]. A remarkable study by Wu and co-workers of Mg (dipnacnac), ((DipNCMe)2CH, Dip = 2,6-diisopropylphenyl), a non-nuclear attractor (NNA), indicated a small (yet significant) decrease in length of a Mg-Mg bond in the complex. Despite the challenges of Mg(I) compounds being very reactive and acting as very efficient reducing agents, the authors were able to show that the central Mg-Mg bond decreases from 2.84 to 2.76 Å, a change of 3%, whilst maintaining NNA behaviour to 1.9 GPa [149]. The compression of metal-containing complexes has also been performed to help determine the nature of metal-ligand interactions. Notably, the compressibility of U-U and U-C-H bonds was studied in the diuranium(III) compound $[UN''2]2(\mu\text{-}\eta6{:}\eta6\text{-}C_6H_6)$ (N″ = $N(SiMe_3)_2$), primarily to study agostic interactions of a low-coordinate U^{III} complex [150]. In this study, which was supported by complementary QTAIM and NBO analyses, one particular U-CH bond (3.022(3) Å at ambient pressure) indicated an agostic interaction but in fact does not become agostic until 3.2 GPa, where the U-CH bond decreased to 2.95(2) Å (Fig. 14).

The following examples are highlighted as they have demonstrated the changes in spectroscopic properties that occur in metal complexes as the coordination environment is compressed. In particular, the compressibility of argentophilic interactions

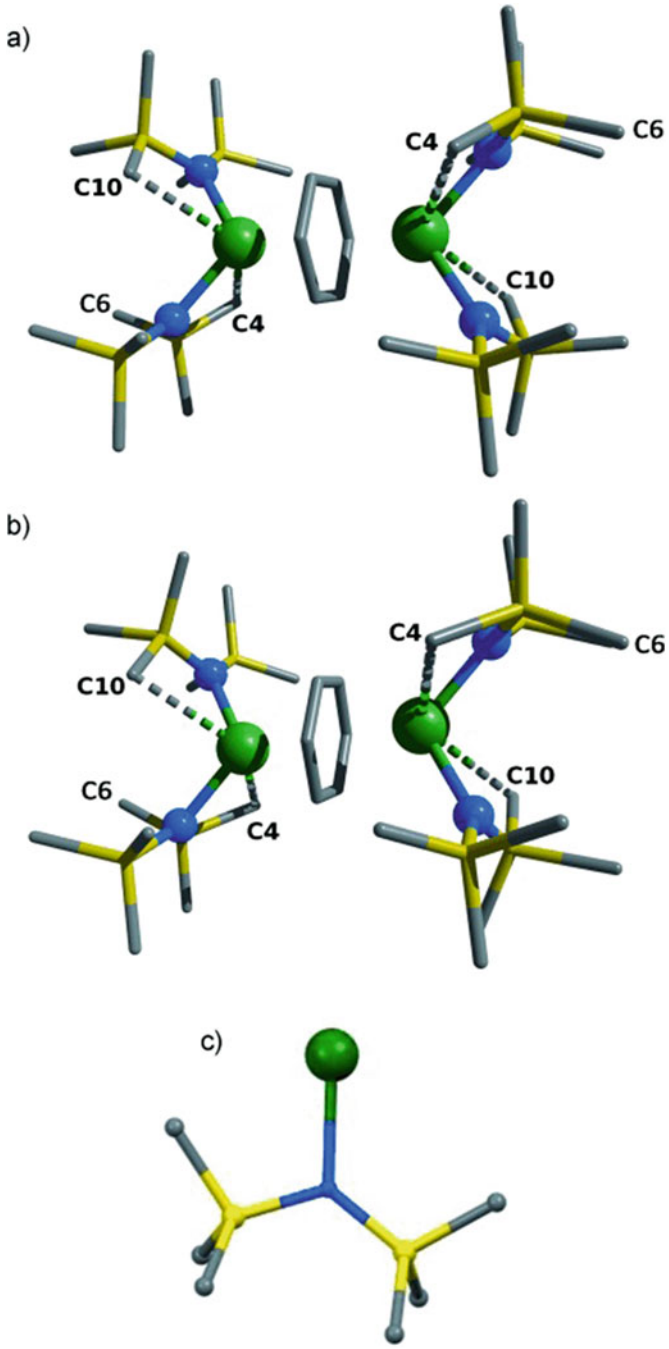


Fig. 14 Structure of $[UN''_2]_2(\mu\text{-}\eta^6\text{:}\eta^6\text{-}C_6H_6)$ ($N'' = N(SiMe_3)_2$) at (**a**) ambient pressure and (**b**) 3.20 GPa. Asymmetry in the N'' ligand is shown in (**c**). H atoms omitted for clarity. Atom colours: green = U; gold = Si; blue = N; grey = C. Figure taken from Angew Chemie Int Ed, 2015, 54:6735–6,739. DOI: (https://doi.org/10.1002/anie.201411250) [150]

has been studied in several Au and Ag complexes. In the tetranuclear Ag^{I}-Cu^{I} complex $Ag_2Cu_2L_4$ (L = 2-diphenylphosphino-3-methylindole) [151], shortening of the Ag-Ag bond was observed, decreasing from 3.042(2) Å under ambient conditions to 2.773(3) Å at 3.47(4) GPa. This is accompanied by a blue-shift to red-shift switch in the absorption and emission spectrum on increasing pressure (between 2 and 3 GPa). The shortening of the Ag-Ag bond was not solely responsible for this, with defects and non-hydrostatic conditions thought to play a role. In the tris(μ_2-3,5-diisopropyl-1,2,4- triazolato-κ^2N^1:N^2)trigold(I) complex, a number of Au...Au intermolecular contacts are also made, with the molecules forming dimers in the solid state [152]. In this study, the compound exhibits four successive phase transitions under hydrostatic pressure, driven by these aurophilic interactions. In this example, the closest Au...Au contacts are between trimeric units where they decrease from 3.4070(5) Å to 3.0273(15) Å, between ambient pressure and 3.31 GPa. A rare observation in this system is the re-emergence of a lower-pressure phase (Form II, 1.69 GPa) at higher pressure (3.3 GPa) after it had undergone transitions to two further phases at 2.18 and 2.70 GPa; this behaviour is referred to as a re-entrant phase transition. Conformational changes in the molecule during the transitions result not only in these shorter Au...Au contacts, but they also correlate with shifts of the luminescence maxima, from a band maximum at 14040 cm^{-1} at 2.40 GPa, decreasing to 13,550 cm^{-1} at 3 GPa.

Spectroscopic changes are not just limited to Au and Ag complexes. A reduction in length of Ni-Ni and Ni-L distances in the complexes bis(3-fluoro-salicylaldoximato)nickel(II) and bis(3-methoxy-salicylaldoximato)nickel(II) has resulted in a blue shift of the UV-Vis leading to an example of piezochromism, with the crystals changing colour from green to red at 5 GPa as the octahedral crystal-field splitting is decreased. In the fluoro system, this equated to a decrease in the Ni...Ni distance from 3.19 to 2.82 Å at 5.4 GPa [153]. The bis(dimethylglyoximato)nickel(II) complex [154] possesses a structure similar to the salicylaldoximato complexes; hence, the changes in the structure (in an equivalent pressure range, to 5.1 GPa) are similar. The glyoximato complex also changes colour but at a lower pressure of 2 GPa. However, the latter complex becomes conducting at high pressure [155] in contrast to the salicyloximato complex that remain non-conducting. Piezochromism, caused by compression of molecular compounds, has also been observed in the photomagnetic compound $[Y(DMF)_4(H_2O)_3(\mu\text{-}CN)Fe(CN)_5]H_2O$ (DMF = dimethylformamide), on increasing pressure to 7.60 GPa, where a reversible yellow-to-red transition is induced via a charge transfer mechanism through the cyanide ligand, though structural evidence for this was lacking as the models could not be refined against the data above 0.7 GPa [156]. A reduction in length of M-M and M-L bonds might seem unsurprising; however, on compression of the compound $Co_2(CO)_6(PPh_3)_2$, the length of the Co-Co bond *increases* from 2.67(1) to 2.72(1) Å due to a change in the geometry of the CO and PPh_3 ligands where they move from a staggered to eclipsed conformation [157].

Conformational flexibility in organic and metal-organic compounds is commonplace within the pressure regime (0.001–10 GPa). In the Pd oxathioether macrocyclic complex $PdCl_2([9]aneS_2O)$, three reported phases of the mononuclear Pd^{II}

complex were studied to 9 GPa. The β-phase was found to undergo a pressure-induced phase transition at 6.8 GPa (designated β'), giving rise to a rearrangement of the macrocycle to give a disordered structure, with each of the three independent molecules in the β' phase showing whole-molecule disorder. This is somewhat unusual, in that disorder is induced on increasing pressure, as often the opposite is observed, with smaller anisotropic displacement parameters and less disorder in a material favouring a reduction in volume [158]. Nevertheless, it is not unheard of, especially in solid-state materials, such as spinels [159].

4.3 Pressure-Induced Bond Formation and Breaking

Although the compression of M-M and M-L bonds is generally observed in metal-containing compounds, examples of bond breaking and bond making are not nearly as common. One extraordinary example was shown in a compression study of the Cu-containing carborane copper(I) m-carborane-9-thiolate (referred to as Cu-S-M9) crystals (where m denotes the meta positions of the carbon atoms in the carborane; see Fig. 15) [160]. On increasing pressure above 8 GPa, elemental Cu is produced in the form of nanoparticles caused by a mechanochemical reduction of the Cu within the carborane from Cu(I) to Cu(0). A reaction which is complete by 12 GPa. Energy dispersive X-ray spectroscopy showed that the nanoclusters were composed exclusively of copper and free of sulphur, whilst the average size of the Cu nanoclusters (≈ 10 nm) was determined by the size of the Cu-S-M9 carborane itself. The use of pressure to polymerise materials and form bonds has been known for some time (see Sect. 3.2.2), but the application of pressure here to form discrete nanoparticles is very unique.

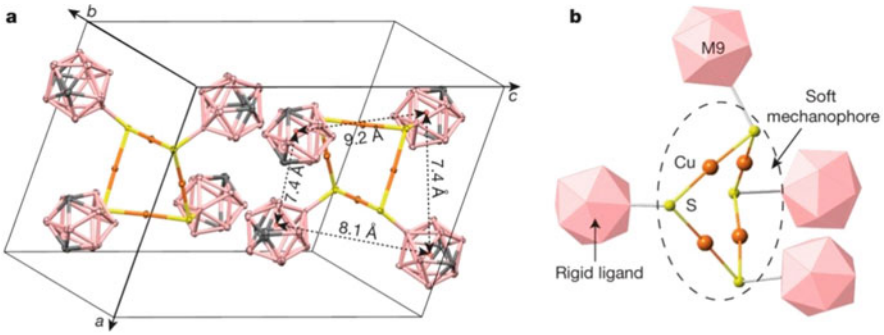


Fig. 15 (**a**) Unit cell of Cu-S-M9. Atoms are represented by their thermal ellipsoids at the 50% probability level. Copper, sulphur, carbon and boron atoms are represented by red, yellow, grey and pink ellipsoids, respectively. Hydrogen atoms and interstitial solvent (toluene) molecules are omitted for clarity. (**b**) Cu-S-M9 molecule showing the Cu_4S_4 mechanophore surrounded by M9 ligands, represented by polyhedra. Figure adapted from Nature, 2018, 554:505–510, DOI: (https://doi.org/10.1038/nature25765) [160]

Nevertheless, bond formation is not always favoured at pressure. In the one-dimensional Cu-(II)-containing coordination polymer $(Cu_2L_2(1\text{-methylpiperazine})_2]n$ (where H_2L = 1,1′-(1,3-phenylene)-bis (4,4-dimethylpentane-1,3-dione)), pressure-induced Cu-N bond breaking/bond forming occurs at a very low pressure of only 0.05 GPa, immediately on loading the crystals inside the DAC. The phase change results in the depolymerisation of the material through the cleavage and formation of axial Cu-N bonds, resulting in the formation of a discrete dinuclear complex [161]. The breaking and subsequent release of coordinated water ligands was also observed in the hydroxo-bridged Cu (II) dimer $[CuF_2(H_2O)_2(pyz)]$ (1, pyz = pyrazine) [162]. In this study, two phase transitions were observed, with the second transition (occurring between 2.85 and 3.3 GPa), resulting in a loss in symmetry and ejection of one water molecule per copper unit from two thirds of the chains, forcing them to dimerise through the F-atoms to fill the vacant coordination sites. This study was performed in parallel with high-pressure EPR, with the main focus of the paper being a Jahn-Teller switching within the material, which we refer to in the next section.

4.4 Functional Materials at Pressure

4.4.1 Molecular Magnetic Materials

Structure/property relationships in a wide range of functional materials are usually established by systematically modifying the chemical structure to produce a number of related compounds, complexes or frameworks, whose physical properties can be measured and observations correlated to the molecular or crystal structure. This approach has been used successfully in molecular magnetic materials such as dinuclear metal complexes, where the geometrical parameters between the metals and associated bridging ligands can affect whether the metals interact ferro- or antiferromagnetically [163]. A more recent approach has been to perform high-pressure crystallographic experiments, in parallel with high-pressure magnetic measurements, to directly measure how changes in the intramolecular geometry and intermolecular interactions affect magnetic behaviour. Some of this behaviour has already been detailed in the review by Moggach et al. [143], though several other examples have been observed since the review was published. One of the main structural concepts that has been explored in molecular magnetic materials is the Jahn-Teller distortion around the metal centre. This is exemplified by the Mn_{12} complex $Mn_{12}O_{12}(O_2CCH_2{}^tBu)_{16}(H_2O)_4{\cdot}CH_2Cl_2{\cdot}MeNO_2$. The complex consists of a cube of $Mn(IV)_2O_4$ surrounded by eight Mn(III) metal sites that are bonded to the ligands. As pressure is applied to this system, the compression of the Jahn-Teller axes on each of the eight Mn(III) centres differs slightly with one of the centres switching the Jahn-Teller axes above 1.5 GPa so that they lie normal to the plane made of the two most compressible directions [164]. This transition was also accompanied by a release of solvent from the crystal into the PTM on compression,

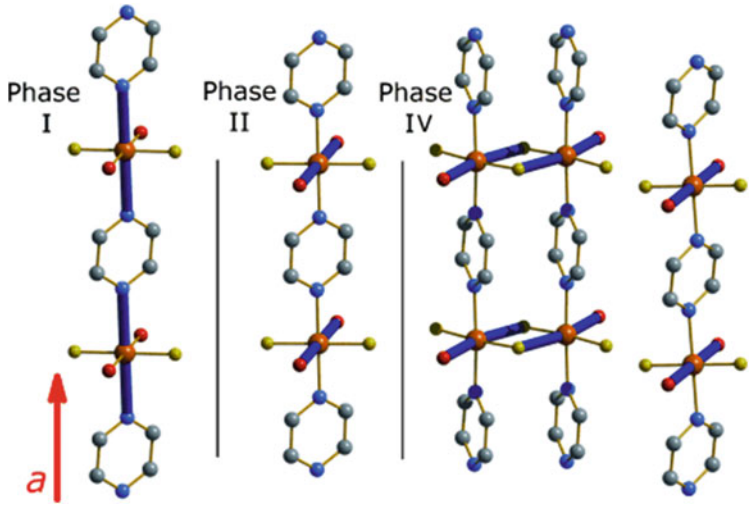


Fig. 16 Disposition of the JT axes (highlighted by the dark blue rods) in the three phases of $CuF_2(H_2O)_2(pyz)$. Colour scheme: Cu = orange, O = red, F = yellow, N = blue, C = grey. H-atoms have been removed for clarity. Figure taken from Angew. Chem. Int. Ed., 2012, Volume: 51, Issue: 30, Pages: 7490–7,494, DOI: (https://doi.org/10.1002/anie.201202367) [162]

and reabsorbed on decompression. The net result in the magnetic properties is the conversion of the complex from a fast-relaxing to slow-relaxing species on increasing pressure to 1.44 GPa. Jahn-Teller switching was also observed in the previously mentioned $CuF_2(H_2O)(pyz)$ (Fig. 16) [162, 165]. In this complex, however, the reorientation of the Jahn-Teller has a profound influence on the magnetic properties, with the Jahn-Teller switching promoting appreciable correlations along the Cu-pyz-Cu chains. This was followed rather elegantly with high-pressure EPR measurements, where the crystals were aligned along their Jahn-Teller axis perpendicular to the plane of rotation. In another Cu(II) chain-polymer complex $[Cu(hfac)_2L]$-II (hfac = hexafluoroacetylacetone, L = 4,4,5,5-tetramethyl-2-(1-methyl-1H-pyrazol-4-yl)-imidazoline-3-oxide-1-oxyl), the Cu centres also undergo a Jahn-Teller switch on increasing pressure; however, in addition to this, a strengthening of the antiferromagnetic exchange occurs between the neighbouring nitroxyl groups, caused by a drastic change in the O…O distance between neighbouring chains [166]. Pressure-induced Jahn-Teller switching behaviour, where elongated M-L bonds switch from being axial to equatorial, seems to be a theme in this class of material, with a couple of examples previously observed [167, 168]; however, the suppression of the Jahn-Teller axis would appear to be more common. For example, in the one-dimensional coordination polymer $[Cu(II)(L\text{-aspartate})(H_2O)_2]$, where a distant carboxylic oxygen located at 2.925(2) Å compresses to 2.883(6) Å, or in the single-molecular magnet $[Mn_6O_2(2\text{-hydroxyphenylpropanone})_6(L)_2(EtOH)_6]$, (L = $-O_2CPh(Me)_2$ or naphthalene carboxylate), where a general suppression of the Jahn-Teller axes is

observed, decreasing in length by 0.075(6) Å for L = $-O_2CPh(Me)_2$ and 0.081(8) Å for L = naphthalene carboxylate, on average [169].

In all of the aforementioned complexes, the twisting or distortion of ligands between metal centres within a molecular magnet results in a change in the magnetic exchange between the metal centres. Another area of interest in the world of molecular magnetism is the study of single-ion magnets (or SIMs). A significant feature required for molecular magnets to be used for information storage is the presence of giant axial magnetic anisotropy. The barrier to magnetic relaxation, which would cause loss of data, is determined by the size of the spin ground state and the size of the axial anisotropy. Here, the focus has been on achieving exquisite control of the coordination environment around a single paramagnetic metal ion, to generate a ligand field that leads to first-order spin-orbit coupling. In the SIM $[Ni(MeDABCO)_2Cl_3]^+$ (MeDABCO = 1-methyl-4-aza-1-azoniabicyclo[2.2.2] octanium), the application of pressure causes the giant axial magnetic anisotropy to decrease as the symmetry about the Ni centre [170]. Here, as in the previous examples, pressure is used to distort the metal centre to determine which parameters control the magnetic anisotropy, rather than modifying the ligands around the metal centre. Here, the Cl-Ni-Cl angles, in particular, were particularly prone to distortion, with two of the three Cl-Ni-Cl angles increasing in size (the largest by $\approx 5°$), whilst the third actually decreased by $\approx 9°$.

4.4.2 Spin-Crossover Complexes

Metal complexes particularly first-row transition metal complexes with a d^4 to d^7 configuration in an octahedral geometry are well-known to undergo transitions from high- to low-spin states or vice versa. This field is dominated by Fe(II)-containing spin-crossover complexes, where changes in spin state can be easily followed by measuring magnetic susceptibility, and even performing Mössbauer spectroscopy, as the change from a paramagnetic to diamagnetic species can be easily followed. Transitions between these two spin states can be induced by temperature, pressure or light. Pressure is particularly useful, as increasing pressure favours the low-spin (smaller volume) spin state which may be inaccessible via any other route. In the complex $[\{Fe(bpp)(NCS)_2\}_2(4,4'\text{-bipy})]\cdot 2MeOH$, (bpp = 2,6-bis(pyrazol-3-yl)pyridine, 4,4′-bipy = 4,4′-bipyridine), increasing pressure above 0.7 GPa induces a spin-crossover transition, which is not thermally accessible. It does this without causing a crystallographic phase transition (i.e. major structural change) in the crystal [171]. Phase transitions involving a change in symmetry, or change in size of the unit cell dimensions, however, are often observed. In the complex $[FeII(bapbpy)(NCS)_2]$ [172] (bapbpy = 6,6-bis(amino-2-pyridyl)-2,2-bipyridine), a 'stepped' first-order transition from the high-spin to low-spin phase was observed upon compression, where on increasing pressure, an intermediate phase (between a fully high-spin and low-spin state) was observed between 0.4 and 1.1 GPa. This phase was characterised by supercell reflections and tripling of the *c*-axis of the unit cell due to the formation of a periodic [HS-LS-LS] structural motif. Interestingly, in

the complex $Fe(3\text{-}OMeSalEen)_2]PF_6$ [173] (H-3-OMeSalEen is a condensation product of 3-methoxysalicylaldehyde and N-ethylethylenediamine), the influence of PTM on the structural changes is observed in this system (cf. Sect. 2.2.4). The three media used were Alcatel 100 (a paraffin-based mineral oil from Alcatel), Fluorinert FC77 (a perfluorocarbon liquid from 3 M) and Daphne 7,373 (a mixture of olefin oligomers from Idemitsu Kosan Global). The change in the PTM results in modification of the spin-crossover curves on decompression, thought to be caused by damage to the crystals on compression and subsequent decompression in these different fluids. Although the structural distortions observed in the aforementioned Fe-based spin-crossover complexes resulted in a distortion of the individual molecules, the spin transition in the molecular complex $[Fe(dpp)_2(NCS)_2]\cdot py$ (dpp = dipyrido[3,2-α:2′3′-*c*]phenazine) gives rise to a 'scissor'-like motion, as coined by the authors, which not only results in a distortion of the individual molecules but also results in NLC behaviour, with the isothermal compressibility $K_a = -30(4)\ TPa^{-1}$ [174].

4.5 Metal Organic Frameworks and Coordination Polymers

According to IUPAC, metal-organic frameworks (or MOFs) are defined as 'coordination networks with organic ligands containing potential voids' [175]. In MOFs, a three-dimensional framework is generated by the reticular assembly of metal ions or clusters and multi-dentate organic bridges into an infinite polymeric network. The ordered nature of the framework often results in the formation of crystals or polycrystalline powders, although amorphous and liquid MOFs have also recently been reported [176, 177]. Porosity is a common but not pervasive feature of MOFs, with void space accounting for 10–70% of the unit cell volume, with the pores ranging from microporous ($d < 2$ nm) to nanoporous ($d = 2$–50 nm). From the diverse array of framework architectures, chemical composition, porosities and functionalities possible, a taxonomy of MOFs has emerged, including zeolitic imidazolate frameworks (ZIFs), Matériaux de l'Institut Lavoisier (MILs) and isoreticular frameworks (IRMOFs). Materials in these classes are some of the most studied at high pressure. Appropriate selection of the PTM is important in the study of porous metal organic framework and in particular with respect to their structural integrity and pressure response. Small, penetrating media, for example (such as gases or liquids), can enter the pores of the framework at elevated pressure. Penetration of methanol into the pores of ZIF-8 at elevated pressure was demonstrated in the work by Moggach et al. [178], in which a solvent-dependent pressure-induced 'gate-opening' rotation of the imidazolate linker was observed. This transition proved pivotal in understanding the maximum amount of guest N_2 uptake under much milder pressure (bars). The phase transition observed at 1.47 GPa in a MeOH: EtOH mixture was established to be the same transition observed on exposing a polycrystalline powder of ZIF-8 to N2 gas at 0.4 bar [3]. Prior to this pressure study, ZIF-8 was actually considered a rather rigid MOF, with other classes of MOFs, such

as MILs, more notable for their flexibility on uptake of guest species [179]. Interestingly, since this transition was published, others have since shown that the transition is also crystallite size dependent, with larger nanosized crystals not undergoing the same transition on exposure to N_2 gas [180]. In 2016, in a study by Hobday et al. [181], the crystallite size dependence was overcome by loading N_2 as a PTM at 2,000 bar, inducing the transition to the 'open' structure. Crystallographic measurements were also accompanied by complementary DFT and Grand Canonical Monte Carlo (GCMC) simulations, which showed that the transition was driven by the formation of a new favourable framework binding site for the N_2 gas molecules inside the pores.

The ability to probe the effect of PTM guest uptake in MOFs and explore guest-dependent behaviour under pressure has resulted in a number of studies for the MOFs HKUST-1 [182, 183], MOF-5 [184], Sc_2BDC_3 [185], ZIF-7 [186] and MIL-47 [187], though the latter study here on MIL-47 is conducted at MPa rather than GPa pressures utilising mercury intrusion experiments. The review by McKellar et al. [50] covers the majority of behaviour on these MOFs; however, a significant body of work has been conducted on the isoreticular UiO-MOFs [188] (Universitetet i Oslo) since 2015.

UiO-MOFs comprise hexanuclear metal nodes, $M_6O_4(OH)_4$ (where M = Zr or Hf) linked by carboxylate bridges. Currently, there are 11 unique native UiO-type structures in the CSD [97]. High-pressure studies have focussed on the isostructural series, UiO-66 [189], UiO-67 [190], UiO-68 [191] and UiO-abdc, which feature 1,4-benzendicarboxylate (BDC),biphenyl-4,4′-dicarboxylate (BPDC), p-terphenyl-4,4′-dicarboxylic acid (TPDC) and azobenzene-4,4′-dicarboxylate ligands, respectively (ABDC). Each crystallises in the cubic space group Fm-3 m and contains a large central octahedral pore that shares faces with eight smaller tetrahedral voids. Thermal and mechanical stability is provided to the framework by the highly coordinated metal ions and the 12-connected paddle-wheel nodes. These connections limit the flexibility of the framework, making the elastic properties of UiO-frameworks amongst the highest reported for porous MOFs (E ~ 20–59 GPa, K ~ 8–65 GPa) [192–195].

The framework resilience of UiO-66 and its amine-functionalised derivative, UiO-66-NH_2, was demonstrated by Yot et al. [192] and Peterson et al. [196, 197], who reported no structural degradation during isotropic compression up to 2 GPa or pelletisation up to 0.17 GPa. Functionalisation with $–NH_2$ groups provided improved structural stability and higher incompressibility to the native framework, increasing its bulk modulus from 17 GPa (UiO-66) to 25 GPa (UiO-66-NH_2). Corresponding behaviour was reported for nitro-functionalised scandium terephthalate, Sc_2BDC_3-NO_2, which resisted amorphisation compared to the parent material [185]. It must be noted that on increasing pressure on MOFs, crystalline-amorphous transitions with a resulting loss in Bragg scattering appear to be frequently observed.

Recently, the mechanical properties of UiO-66, UiO-67 and UiO-abdc have been examined by compression in a hydraulic piston press, TEM punch experiments and first-principles calculations. Under anisotropic pressure, UiO-66 loses crystallinity at only 0.4 GPa. Similar behaviour was predicted under hydrostatic compression from

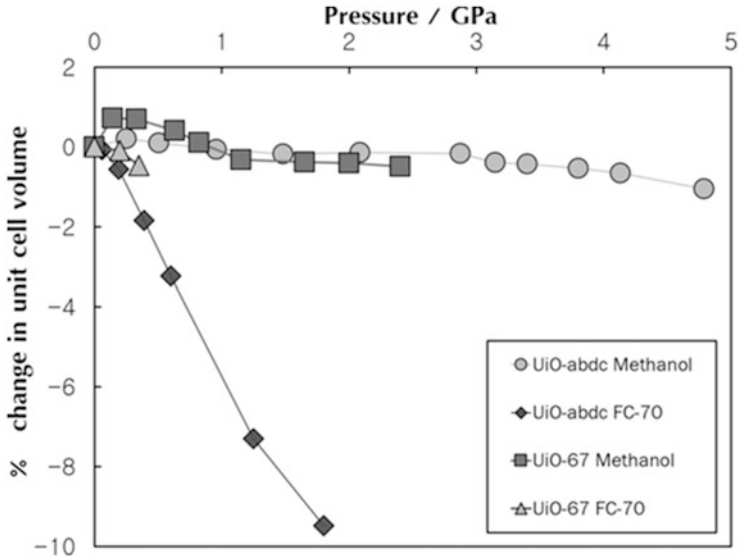


Fig. 17 Graph of percentage change in volume vs. pressure (GPa) for UiO-abdc in methanol (circles), FC-70 (diamonds), and UiO-67 in methanol (squares), FC-70 (triangles). Figure taken from Angew. Chem. Int. Ed., 2016, Volume: 128, Issue: 7, Pages: 2447–2451, DOI: (https://doi.org/10.1002/ange.201509352) [190]

DFT simulations [190]. Introduction of defects into the structure has also provided additional control over the compressibility and pressure at which amorphisation of the MOF occurs. A method to create defects in these materials has been the addition of controlled equivalents of trifluoroacetic acid modulator during preparation of UiO-66 to introduce missing-ligand defects in the framework, lowering the average coordination number of the $Zr_6O_4(OH)_4$ nodes [194]. Decreasing the connection of the net from 12c (non-defected) to 9c (defected) lowers the amorphisation onset pressure from 0.4 to 0.3 GPa and decreases the bulk modulus of the material (K_0) from ~26 GPa to ~12 GPa (K_0 was calculated with 8 data points up to 0.175 GPa). Further decrease of the net connection to 8c increased the bulk modulus to ~14 GPa due to the suspected transition of the network topology.

As is typical for porous materials, guest-loaded UiO-type frameworks are resistive to compression. Pressure-induced pore filling of UiO-67 and UiO-*abdc* with methanol increases the bulk moduli of the materials up to 30-fold that of the native framework, going from 13.4 GPa to 174 GPa for UiO-67 and 16.8 GPa to 580 GPa for UiO-abdc from theoretical simulations and nanoindentation measurements. Maximum occupation of the pores was reached by ~1.2 GPa, signalled by no further volumetric increase of the unit cell (Fig. 17). Notably, both materials were unaffected by further compression up to 2.4 GPa (UiO-67) or 4.8 GPa (UiO-abdc). Slight compression of the unit cell is typical for MOFs with fully occupied pores; the total resistance to compression of these UiO-frameworks is therefore unusual.

5 Concluding Remarks

We hope that this chapter has provided a deeper insight into the methods and practices that have been used to explore the solid-state forms of organic and metal-organic materials under high-pressure conditions. We wanted to provide a framework of practice that could be used as a reference text for those entering into the area of high-pressure research. We have tried to provide a balanced view of the science conducted in the areas of organic and metal-organic solid-state chemistry, but inevitably there will be studies that have been overlooked due to the sheer volume of data and articles in the literature; if we have missed a few, we can only apologise. As you will have read, the area of high pressure on molecular materials is vast even considering that there are many spectroscopic studies of materials at high pressure that we have not discussed here in detail. The future for high-pressure science on molecular systems is bright, and with advances of techniques, only the imagination is the kinetic barrier to discovery in the twenty-first century and beyond.

References

1. Oswald IDH, Lennie AR, Pulham CR, Shankland K (2010) High-pressure structural studies of the pharmaceutical, chlorothiazide. CrystEngComm 12:2533
2. Moggach SA, Parsons S, Wood PA (2008) High-pressure polymorphism in amino acids. Crystallogr Rev 14:143–184
3. Fairen-Jimenez D, Moggach SA, Wharmby MT, Wright PA, Parsons S, Düren T (2011) Opening the gate: framework flexibility in ZIF-8 explored by experiments and simulations. J Am Chem Soc 133:8900–8902
4. Yoo CS, Cynn H (1999) Equation of state, phase transition, decomposition of β-HMX (octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine) at high pressures. J Chem Phys 111:10229–10235
5. Millar DIA, Marshall WG, Oswald IDH, Pulham CR (2010) High-pressure structural studies of energetic materials. Crystallogr Rev 16:115–132
6. Neumann MA, Van De Streek J, Fabbiani FPA, Hidber P, Grassmann O (2015) Combined crystal structure prediction and high-pressure crystallization in rational pharmaceutical polymorph screening. Nat Commun 6:7793
7. Bhardwaj RM, McMahon JA, Nyman J, Price LS, Konar S, Oswald IDH, Pulham CR, Price SL, Reutzel-Edens SM (2019) A prolific solvate former, galunisertib, under the pressure of crystal structure prediction, produces ten diverse polymorphs. J Am Chem Soc 141:13887–13897
8. Boldyreva EV (2008) High-pressure diffraction studies of molecular organic solids. A personal view. Acta Crystallogr Sect A Found Crystallogr 64:218–231
9. Fabbiani FPA, Pulham CR (2006) High-pressure studies of pharmaceutical compounds and energetic materials. Chem Soc Rev 35:932–942
10. Katrusiak A (2008) High-pressure crystallography. Acta Cryst 64:135–148
11. Miletich R, Allan DR, Kuhs WF (2000) High-pressure single-crystal techniques. Rev Mineral Geochem 41:445–519
12. Sidorov NV, Krasyukov YN, Mukhtarova ÉI, Shkrabo DM (2000) Phase transition and domain structure in iodoform crystals. Crystallogr Rep 45:288–293

13. Dziubek KF, Katrusiak A (2008) Polar symmetry in new high-pressure phases of chloroform and bromoform. J Phys Chem B 112:12001–12009
14. Allan DR, Clark SJ (1999) Comparison of the high-pressure and low-temperature structures of ethanol and acetic acid. Phys Rev B Condens Matter Mater Phys 60:6328–6334
15. Allan DR, Clark SJ, Brugmans MJP, Ackland GJ, Vos WL (1998) Structure of crystalline methanol at high pressure. Phys Rev B 58:R11809–R11812
16. Ridout J, Probert MR (2014) Low-temperature and high-pressure polymorphs of isopropyl alcohol. CrystEngComm 16:7397–7400
17. Bull CL, Funnell NP, Pulham CR, Marshall WG, Allan DR (2017) A new high-pressure polymorph of phosphoric acid. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 73:1068–1074
18. Bridgman PW (1914) Change of phase under pressure. I. The phase diagram of eleven substances with especial reference to the melting curve. Phys Rev 3:126–141
19. Dziubek K (2014) Complementing diffraction data with volumetric measurements. Zeitschrift für Krist Cryst Mater 229:129–134
20. Marciniak J, Bakowicz J, Dobrowolski MA, Dziubek KF, Kaźmierczak M, Paliwoda D, Rajewski KW, Sobczak S, Stachowicz M, Katrusiak A (2016) Most frequent organic interactions compressed in toluene. Cryst Growth Des 16:1435–1441
21. Oswald IDH, Crichton WA (2009) Structural similarities of 2-chlorophenol and 2-methylphenol. CrystEngComm 11:463–469
22. Funnell NP, Dawson A, Marshall WG, Parsons S (2013) Destabilisation of hydrogen bonding and the phase stability of aniline at high pressure. CrystEngComm 15:1047–1060
23. Oswald IDH, Allan DR, Motherwell WDS, Parsons S (2005) Structures of the monofluoro- and monochlorophenols at low temperature and high pressure. Acta Crystallogr Sect B Struct Sci 61:69–79
24. Paliwoda D, Dziubek KF, Katrusiak A (2012) Imidazole hidden polar phase. Cryst Growth Des 12:4302–4305
25. Delori A, Hutchison IB, Bull CL, Funnell NP, Urquhart AJ, Oswald IDH (2018) Reaction of acetylenedicarboxylic acid made easy: high-pressure route for polymerization. Cryst Growth Des 18:1425–1431
26. Allan DR, Clark SJ, Ibberson RM, Parsons S, Pulham CR, Sawyer L (1999) The influence of pressure and temperature on the crystal structure of acetone. Chem Commun:751–752
27. Vos WL, Finger LW, Hemley RJ, Hu JZ, Mao HK, Schouten JA (1992) A high-pressure van der Waals compound in solid nitrogen-helium mixtures. Nature 358:46–48
28. Vos WL, Finger LW, Hemley RJ, Mao HK (1993) Novel H2-H2O clathrates at high pressures. Phys Rev Lett 71:3150–3153
29. Weck G, Loubeyre P, LeToullec R, LeToullec R (2002) Observation of structural transformations in metal oxygen. Phys Rev Lett 88:4
30. Lundegaard LF, Weck G, McMahon MI, Desgreniers S, Loubeyre P (2006) Observation of an O8 molecular lattice in the epsilon phase of solid oxygen. Nature 443:201–204
31. Oswald IDH, Urquhart AJ (2011) Polymorphism and polymerisation of acrylic and methacrylic acid at high pressure. CrystEngComm 13:4503–4507
32. Murli C, Song Y (2010) Pressure-induced polymerization of acrylic acid: a Raman spectroscopic study. J Phys Chem B 114:9744–9750
33. Johnston BF, Marshall WG, Parsons S, Urquhart AJ, Oswald IDH (2014) Investigation of acrylic acid at high pressure using neutron diffraction. J Phys Chem B 118:4044–4051
34. Marshall WG, Urquhart AJ, Oswald IDH (2015) Investigation of methacrylic acid at high pressure using neutron diffraction. J Phys Chem B 119:12147–12154
35. Marciniak J, Katrusiak A (2017) Direct and inverse relations between temperature and pressure effects in crystals: a case study on o-xylene. J Phys Chem C 121:22303–22309
36. Haisa M, Kashino S, Kawai R, Maeda H (1976) The monoclinic form of p-hydroxyacetanilide. Acta Crystallogr Sect B Struct Crystallogr Cryst Chem 32:1283–1285

37. Haisa M, Kashino S, Maeda H (1974) The orthorhombic form of p-hydroxyacetanilide. Acta Crystallogr Sect B Struct Crystallogr Cryst Chem 30:2510–2512
38. Perlovich GL, Volkova TV, Bauer-Brandl A (2007) Polymorphism of paracetamol. J Therm Anal Calorim 89:767–774
39. Nichols G, Frampton CS (1998) Physicochemical characterization of the orthorhombic polymorph of paracetamol crystallized from solution. J Pharm Sci 87:684–693
40. Boldyreva EV, Shakhtshneider TP, Ahsbahs H, Sowa H, Uchtmann H (2002) Effect of high pressure on the polymorphs of paracetamol. J Therm Anal Calorim 68:437–452
41. Fabbiani FPA, Allan DR, Dawson A, David WIF, McGregor PA, Oswald IDH, Parsons S, Pulham CR (2003) Pressure-induced formation of a solvate of paracetamol. Chem Commun (Camb) 3:3004–3005
42. Fabbiani FPA, Allan DR, David WIF, Moggach SA, Parsons S, Pulham CR (2004) High-pressure recrystallisation - a route to new polymorphs and solvates. CrystEngComm 6:504–511
43. Oswald IDH, Chataigner I, Elphick S, Fabbiani FPA, Lennie AR, Maddaluno J, Marshall WG, Prior TJ, Pulham CR, Smith RI (2009) Putting pressure on elusive polymorphs and solvates. CrystEngComm 11:359–366
44. Fabbiani FPA, Allan DR, David WIF, Davidson AJ, Lennie AR, Parsons S, Pulham CR, Warren JE (2007) High-pressure studies of pharmaceuticals: an exploration of the behavior of piracetam. Cryst Growth Des 7:1115–1124
45. Fabbiani FPA, Buth G, Levendis DC, Cruz-Cabeza AJ (2014) Pharmaceutical hydrates under ambient conditions from high-pressure seeds: a case study of GABA monohydrate. Chem Commun (Camb) 50:1817–1819
46. Ledru J, Imrie CT, Pulham CR, Céolin R, Hutchinson JM (2007) High pressure differential scanning calorimetry investigations on the pressure dependence of the melting of paracetamol polymorphs I and II. J Pharm Sci 96:2784–2794
47. Ward MR, Oswald IDH (2019) Antisolvent addition at extreme conditions. CrystEngComm 21:4437–4443
48. Fabbiani FPA, Allan DR, Parsons S, Pulham CR (2005) An exploration of the polymorphism of piracetam using high pressure. CrystEngComm 7:179
49. Collings IE, Hanfland M, Collings IE, Hanfland M (2019) Packing rearrangements in 4-hydroxycyanobenzene under pressure. Molecules 24:1759
50. McKellar SC, Moggach SA (2015) Structural studies of metal-organic frameworks under high pressure. Acta Crystallogr Sect B 71:587–607
51. Eikeland E, Thomsen MK, Madsen SR, Overgaard J, Spackman MA, Iversen BB (2016) Structural collapse of the hydroquinone-formic acid clathrate: a pressure-medium-dependent phase transition. Chem Eur J 22:4061–4069
52. Eikeland E, Thomsen MK, Overgaard J, Spackman MA, Iversen BB (2017) Intermolecular interaction energies in hydroquinone clathrates at high pressure. Cryst Growth Des 17:3834–3846
53. Merrill L, Bassett WA (1974) Miniature diamond anvil pressure cell for single crystal x-ray diffraction studies. Rev Sci Instrum 45:290–294
54. Boehler R, De Hantsetters K (2004) New anvil designs in diamond-cells. High Press Res 24:391–396
55. Moggach SA, Allan DR, Parsons S, Warren JE (2008) Incorporation of a new design of backing seat and anvil in a Merrill–Bassett diamond anvil cell. J Appl Crystallogr 41:249–251
56. Yamanaka T, Fukuda T, Hattori T, Sumiya H (2001) New diamond anvil cell for single-crystal analysis. Rev Sci Instrum 72:1458
57. Kenichi T, Satoshi N (2003) Performance of a synthetic diamond backing-plate for the diamond-anvil cell at ultrahigh pressures. Rev Sci Instrum 74:3017
58. Ahsbahs H (2004) New pressure cell for single-crystal X-ray investigations on diffractometers with area detectors. Zeitschrift für Krist Cryst Mater 219:305–308

59. Fabbiani FPA, Buth G, Dittrich B, Sowa H (2010) Pressure-induced structural changes in wet vitamin B12. CrystEngComm 12:2541
60. Grzechnik A, Meven M, Friese K (2018) Single-crystal neutron diffraction in diamond anvil cells with hot neutrons. J Appl Crystallogr 51:351–356
61. Haberl B, Dissanayake S, Wu Y et al (2018) Next-generation diamond cell and applications to single-crystal neutron diffraction. Rev Sci Instrum 89:092902
62. Binns J, Kamenev KV, McIntyre GJ, Moggach SA, Parsons S (2016) Use of a miniature diamond-anvil cell in high-pressure single-crystal neutron Laue diffraction. IUCrJ 3:168–179
63. Akella J, Kennedy GC (1971) Phase diagram of benzene to 35 kbar. J Chem Phys 55:793
64. Lomberget T, Chataigner I, Bouyssi D, Maddaluno J, Balme G (2004) Diels–Alder cycloadditions of functionalized (Z)-1-benzylidene-2-methylene cyclohexanes: the beneficial effect of high pressure. Tetrahedron Lett 45:3437–3441
65. Hutchison IB, Bull CL, Marshall WG, Parsons S, Urquhart AJ, Oswald IDH (2017) Compression of glycolide-h4 to 6GPa. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 73:1151–1157
66. Ward MR, Younis S, Cruz-Cabeza AJ, Bull CL, Funnell NP, Oswald IDH (2019) Discovery and recovery of delta p -aminobenzoic acid. CrystEngComm 21:2058–2066
67. Zieliński W, Katrusiak A (2015) Hydrate smaller than the anhydrate. CrystEngComm 17:5468–5473
68. Millar DIA, Oswald IDH, Barry C, Francis DJ, Marshall WG, Pulham CR, Cumming AS (2010) Pressure-cooking of explosives - the crystal structure of ε-RDX as determined by X-ray and neutron diffraction. Chem Commun 46:5662–5664
69. Wang W, Fortes AD, Dobson DP, Howard CM, Bowles J, Hughes NJ, Wood IG (2018) Investigation of high-pressure planetary ices by cryo-recovery. II. High-pressure apparatus, examples and a new high-pressure phase of MgSO 4 ·5H 2 O. J Appl Crystallogr 51:692–705
70. Hutchison IB, Delori A, Wang X, Kamenev KV, Urquhart AJ, Oswald IDH (2015) Polymorphism of a polymer precursor: metastable glycolide polymorph recovered via large scale high-pressure experiments. CrystEngComm 17:1778–1782
71. Wood IG, Fortes AD, Dobson DP, Wang W, Pajdzik L, Cosier J (2018) Investigation of high-pressure planetary ices by cryo-recovery. I. An apparatus for X-ray powder diffraction from 40 to 315 K, allowing `cold loading' of samples. J Appl Crystallogr 51:685–691
72. Roszak K, Katrusiek A, Katrusiak A (2016) High-pressure preference for the low Z′ polymorph of a molecular crystal. Cryst Growth Des 16:3947–3953
73. Yan TT, Xi DY, Wang JH, Fan XF, Wan Y, Zhang LX, Wang K (2019) High-pressure-induced phase transition in cinchomeronic acid polycrystalline form-I. Chinese Phys B 28:16104
74. Abbas N, Oswald IDH, Pulham CR (2017) Accessing mefenamic acid form II through high-pressure recrystallisation. Pharmaceutics 9:16
75. Cosier J, Glazer AM (1986) A nitrogen-gas-stream cryostat for general X-ray diffraction studies. J Appl Crystallogr 19:105–107
76. Brock CP, Duncan LL (1994) Anomalous space-group frequencies for monoalcohols CnHmOH. Chem Mater 6:1307–1312
77. McGregor PA, Allan DR, Parsons S, Pulham CR (2005) The low-temperature and high-pressure crystal structures of cyclobutanol (C4H7OH). Acta Crystallogr Sect B Struct Sci 61:449–454
78. Moggach SA, Allan DR, Lozano-Casal P, Parsons S (2005) High-pressure polymorphism of cyclopentanol (C5H10O): the structure of cyclopentanol phase-V at 1.5 GPa. J Synchrotron Radiat 12:590–597
79. Mcgregor PA, Allan DR, Parsons S, Clark SJ (2006) Structural science hexamer formation in tertiary butyl alcohol (2-methyl-2-propanol, C 4 H 10 O). Res Pap Acta Cryst 62:599–605
80. Oswald IDH, Allan DR, Day GM, Motherwell WDS, Parsons S (2005) Realizing predicted crystal structures at extreme conditions: the low-temperature and high-pressure crystal structures of 2-chlorophenol and 4-fluorophenol. Cryst Growth Des 5:1055–1071

81. Allan DR, Clark SJ, Dawson A, McGregor PA, Parsons S (2002) Pressure-induced polymorphism in phenol. Acta Crystallogr Sect B Struct Sci 58:1018–1024
82. Podsiadło M, Patyk E, Katrusiak A (2012) Chiral aggregation hierarchy in high-pressure resolved 2-butanol and 2,3-butanediol. CrystEngComm 14:6419
83. Cavallo G, Metrangolo P, Milani R, Pilati T, Priimagi A, Resnati G, Terraneo G (2016) The halogen bond. Chem Rev 116:2478–2601
84. Bertolotti F, Curetti N, Benna P, Gervasio G (2013) The effects of P–T changes on intermolecular interactions in crystal structure of iodoform. J Mol Struct 1041:106–112
85. Ridout J, Probert MR (2013) High-pressure and low-temperature polymorphism in C-H···F-C hydrogen bonded monofluorotoluenes. Cryst Growth Des 13:1943–1948
86. Anioła M, Kwas K, Cai W, Katrusiak A (2016) High-pressure crystallizations of meta-dichlorobenzene and dibromobenzene and their solid solutions scheme 1. Atomic labels in the symmetry-independent molecules (labeled A and B) of. Cryst Growth Des 16:6304–6309
87. Giordano N, Afanasjevs S, Beavers CM et al (2019) The effect of pressure on halogen bonding in 4-iodobenzonitrile. Molecules 24:2018
88. Enkelmann V, Wegner G, Novak K, Wagener KB (1993) Single-crystal-to-single-crystal photodimerization of cinnamic acid. J Am Chem Soc 115:10390–10391
89. Fernandes MA, Levendis DC (2016) Photodimerisation of the α′-polymorph of ortho-ethoxy-trans-cinnamic acid occurs via a two-stage mechanism at 343 K yielding 100% α-truxillic acid. CrystEngComm 18:7363–7376
90. Galica T, Bąkowicz J, Konieczny K, Turowska-Tyrk I (2018) Structural transformations in crystals induced by radiation and pressure. Part 6. The reactivity of difluorocinnamic acids under ambient and high pressures – comparative studies. Cryst Growth Des 18:1636–1644
91. Galica T, Bąkowicz J, Konieczny K, Turowska-Tyrk I (2016) Structural transformations in crystals induced by radiation and pressure. Part 5. The influence of pressure on the course of the photochemical reaction in crystals of 2,6-difluorocinnamic acid. CrystEngComm 18:8871–8879
92. Kaupp G (2003) Solid-state molecular syntheses: complete reactions without auxiliaries based on the new solid-state mechanism. CrystEngComm 5:117
93. Fanetti S, Nobrega MM, Dziubek K, Citroni M, Sella A, McMillan PF, Hanfland M, Bini R (2019) Structure and reactivity of 2,4,6-tricyano-1,3,5-triazine under high-pressure conditions. CrystEngComm 21:4493–4500
94. Wilhelm C, Boyd SA, Chawda S et al (2008) Pressure-induced polymerization of diiodobutadiyne in assembled cocrystals. J Am Chem Soc 130:4415–4420
95. Jin HJ, Plonka AM, Parise JB, Goroff NS (2013) Pressure induced topochemical polymerization of diiodobutadiyne: a single-crystal-to-single-crystal transformation. CrystEngComm 15:3106–3110
96. Poręba T, Ernst M, Zimmer D, Macchi P, Casati N (2019) Pressure-induced polymerization and electrical conductivity of a polyiodide. Angew Chem 131:6697–6701
97. Groom CR, Bruno IJ, Lightfoot MP, Ward SC (2016) The Cambridge structural database. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater B72:171–179
98. Freire PTC (2010) Pressure-induced phase transitions in crystalline amino acids. Raman spectroscopy and x-ray diffraction. NATO Sci Peace Secur Ser B Phys Biophys:559–572
99. Dawson A, Allan DR, Belmonte SA, Clark SJ, David WIF, McGregor PA, Parsons S, Pulham CR, Sawyer L (2005) Effect of high pressure on the crystal structures of polymorphs of glycine. Cryst Growth Des 5:1415–1427
100. Moggach SA, Marshall WG, Rogers DM, Parsons S (2015) How focussing on hydrogen bonding interactions in amino acids can miss the bigger picture: a high-pressure neutron powder diffraction study of ε-glycine. CrystEngComm 17:5315–5328
101. Boldyreva EV, Ivashevskaya SN, Sowa H, Ahsbahs H, Weber HP (2005) Effect of hydrostatic pressure on the γ-polymorph of glycine 1. A polymorphic transition into a new δ-form. Zeitschrift fur Krist 220:50–57

102. Bull CL, Flowitt-Hill G, de Gironcoli S, Küçükbenli E, Parsons S, Pham CH, Playford HY, Tucker MG (2017) ζ-Glycine: insight into the mechanism of a polymorphic phase transition. IUCrJ 4:569–574
103. Tumanov NA, Boldyreva EV, Kolesov BA, Kurnosov AV, Quesada Cabrera R (2010) Pressure-induced phase transitions in L-alanine, revisited. Acta Crystallogr Sect B Struct Sci 66:458–471
104. Funnell NP, Marshall WG, Parsons S (2011) Alanine at 13.6 GPa and its pressure-induced amorphisation at 15 GPa. CrystEngComm 13:5841–5848
105. Funnell NP, Dawson A, Francis D, Lennie AR, Marshall WG, Moggach SA, Warren JE, Parsons S (2010) The effect of pressure on the crystal structure of l-alanine. CrystEngComm 12:2573
106. Tumanov NA, Boldyreva EV (2012) X-ray diffraction and Raman study of DL-alanine at high pressure: revision of phase transitions. Acta Crystallogr Sect B Struct Sci 68:412–423
107. Moggach SA, Allan DR, Clark SJ, Gutmann MJ, Parsons S, Pulham CR, Sawyer L (2006) High-pressure polymorphism in L-cysteine: the crystal structures of L-cysteine-III and L-cysteine-IV. Acta Crystallogr Sect B Struct Sci 62:296–309
108. Minkov VS, Drebushchak VA, Ogienko AG, Boldyreva EV (2011) Decreasing particle size helps to preserve metastable polymorphs. A case study of dl-cysteine. CrystEngComm 13:4417
109. Moggach SA, Allan DR, Morrison CA, Parsons S, Sawyer L (2005) Effect of pressure on the crystal structure of l-serine-I and the crystal structure of l-serine-II at 5.4 GPa. Acta Crystallogr Sect B 61:58–68
110. Wood PA, Francis D, Marshall WG, Moggach SA, Parsons S, Pidcock E, Rohl AL (2008) A study of the high-pressure polymorphs of L-serine using ab initio structures and PIXEL calculations. CrystEngComm 10:1154–1166
111. Moggach SA, Marshall WG, Parsons S (2006) High-pressure neutron diffraction study of L-serine-I and L-serine-II, and the structure of L-serine-III at 8.1 GPa. Acta Crystallogr Sect B Struct Sci 62:815–825
112. Fisch M, Lanza A, Boldyreva E, Macchi P, Casati N (2015) Kinetic control of high-pressure solid-state phase transitions: a case study on L-serine. J Phys Chem C 119:18611–18617
113. Johnstone RDL, Francis D, Lennie AR, Marshall WG, Moggach SA, Parsons S, Pidcock E, Warren JE (2008) High-pressure polymorphism in L-serine monohydrate: identification of driving forces in high pressure phase transitions and possible implications for pressure-induced protein denaturation. CrystEngComm 10:1758–1769
114. Zakharov BA, Kolesov BA, Boldyreva EV (2012) Effect of pressure on crystalline l- and dl-serine: revisited by a combined single-crystal X-ray diffraction at a laboratory source and polarized Raman spectroscopy study. Acta Crystallogr Sect B Struct Sci 68:275–286
115. Giordano N, Beavers CM, Kamenev KV, Marshall WG, Moggach SA, Patterson SD, Teat SJ, Warren JE, Wood PA, Parsons S (2019) High-pressure polymorphism in L-threonine between ambient pressure and 22 GPa. CrystEngComm 21:4444–4456
116. Lozano-Casal P, Allan DR, Parsons S (2008) High-pressure structural study of L-α-glutamine and the use of Hirshfeld surfaces and graph-set notation to investigate the hydrogen bonding present in the structure up to 4.9 GPa. Acta Crystallogr Sect B Struct Sci 64:466–475
117. Moggach SA, Allan DR, Parsons S, Sawyer L, Warren JE (2005) The effect of pressure on the crystal structure of hexagonal L-cystine. J Synchrotron Radiat 12:598–607
118. Davidson AJ, Oswald IDH, Francis DJ, Lennie AR, Marshall WG, Millar DIA, Pulham CR, Warren JE, Cumming AS (2008) Explosives under pressure - the crystal structure of gamma-RDX as determined by high-pressure X-ray and neutron diffraction. CrystEngComm 10:162–165
119. Sheldrick GM, IUCr (2015) *SHELXT* – integrated space-group and crystal-structure determination. Acta Crystallogr Sect A Found Adv 71:3–8
120. Drebushchak TN, Drebushchak VA, Boldyreva EV (2011) Solid-state transformations in the β-form of chlorpropamide on cooling to 100 K. Acta Crystallogr Sect B Struct Sci 67:163–176

121. Drebushchak TN, Drebushchak VA, Pankrushina NA, Boldyreva EV (2016) Single-crystal to single-crystal conformational polymorphic transformation in tolbutamide at 313 K. Relation to other polymorphic transformations in tolbutamide and chlorpropamide. CrystEngComm 18:5736–5743
122. Drebushchak VA, Drebushchak TN, Chukanov NV, Boldyreva EV (2008) Transitions among five polymorphs of chlorpropamide near the melting point. J Therm Anal Calorim 93:343–351
123. Boldyreva EV, Dmitriev V, Hancock BC (2006) Effect of pressure up to 5.5 GPa on dry powder samples of chlorpropamide form-A. Int J Pharm 327:51–57
124. Seryotkin YV, Drebushchak TN, Boldyreva EV (2013) A high-pressure polymorph of chlorpropamide formed on hydrostatic compression of the [alpha]-form in saturated ethanol solution. Acta Crystallogr Sect B 69:77–85
125. Zakharov BA, Goryainov SV, Boldyreva EV (2016) Unusual seeding effect in the liquid-assisted high-pressure polymorphism of chlorpropamide. CrystEngComm 18:5423–5428
126. Zakharov BA, Seryotkin YV, Tumanov NA, Paliwoda DD, Hanfland M, Kurnosov AV, Boldyreva EV (2016) The role of fluids in high-pressure polymorphism of drugs: different behaviour of β-chlorpropamide in different inert gas and liquid media. RSC Adv 6:92629–92637
127. Boldyreva EV, Arkhipov SG, Drebushchak TN et al (2015) Isoenergetic polymorphism: the puzzle of tolazamide as a case study. Chem - A Eur J 21:15395–15404
128. Fedorov AY, Rychkov DA, Losev EA, Zakharov BA, Stare J, Boldyreva EV (2017) Effect of pressure on two polymorphs of tolazamide: why no interconversion? CrystEngComm 19:2243–2252
129. Connor LE, Delori A, Hutchison IB, Nic Daeid N, Sutcliffe OB, Oswald IDH (2015) The ecstasy and the agony; compression studies of 3,4-methylenedioxymethamphetamine (MDMA). Acta Crystallogr Sect B Struct Sci Cryst Eng Mater 71:3–9
130. Derollez P, Dudognon E, Affouard F, Danède F, Correia NT, Descamps M (2010) Ab initio structure determination of phase II of racemic ibuprofen by X-ray powder diffraction. Acta Crystallogr Sect B Struct Sci 66:76–80
131. Freer AA, Bunyan JM, Shankland N, Sheen DB (1993) Structure of (S)-(+)-ibuprofen. Acta Crystallogr Sect C Cryst Struct Commun 49:1378–1380
132. Ostrowska K, Kropidowska M, Katrusiak A (2015) High-pressure crystallization and structural transformations in compressed R, S -ibuprofen. Cryst Growth Des 15:1512–1517
133. Rietveld IB, Barrio M, Tamarit J-L, Do B, Céolin R (2011) Enantiomer resolution by pressure increase: inferences from experimental and topological results for the binary enantiomer system (*R*)- and (*S*)-mandelic acid. J Phys Chem B 115:14698–14703
134. Cai W, Marciniak J, Andrzejewski M, Katrusiak A (2013) Pressure effect on D, L-mandelic acid racemate crystallization. J Phys Chem C 117(14):7279–7285
135. Marciniak J, Andrzejewski M, Cai W, Katrusiak A (2014) Wallach's rule enforced by pressure in mandelic acid. J Phys Chem C 118(8):4309–4313
136. Connor LE, Vassileiou AD, Halbert GW, Johnston BF, Oswald IDH (2019) Structural investigation and compression of a co-crystal of indomethacin and saccharin. CrystEngComm 21:4465–4472
137. Caira MR, Alkhamis KA, Obaidat RM (2004) Preparation and crystal characterization of a polymorph, a monohydrate, and an ethyl acetate solvate of the antifungal fluconazole. J Pharm Sci 93:601–611
138. Gorkovenko EA, Kichanov SE, Kozlenko DP, Belushkin AV, Wąsicki J, Nawrocik W, Mielcarek J, Dubrovinsky LS, Lathe C, Savenko BN (2015) The pressure-induced polymorphic transformations in fluconazole. J Pharm Sci 104:4164–4169
139. Hutchison IB, Bull CL, Marshall WG, Urquhart AJ, Oswald IDH (2019) Pressure-induced polymorphism of caprolactam: a neutron diffraction study. Molecules 24:2174
140. Oswald IDH, Pulham CR (2008) Co-crystallisation at high pressure - an additional tool for the preparation and study of co-crystals. CrystEngComm 10:1114–1116

141. Satthaphut N, Sutcliffe OB, Oswald IDH (2014) Putting the squeeze on mephedrone hydrogen sulfate. Zeitschrift für Krist Cryst Mater 229:101–111
142. Cairns AB, Catafesta J, Levelut C, Rouquette J, van der Lee A, Peters L, Thompson AL, Dmitriev V, Haines J, Goodwin AL (2013) Giant negative linear compressibility in zinc dicyanoaurate. Nat Mater 12:212–216
143. Moggach SA, Parsons S (2009) High pressure crystallography of inorganic and organometallic complexes. Spectrosc Prop Inorg Organomet Compd 40:324–354
144. Tidey JP, Wong HLS, Schröder M, Blake AJ (2014) Structural chemistry of metal coordination complexes at high pressure. Coord Chem Rev 277–278:187–207
145. Gütlich P, Ksenofontov V, Gaspar AB (2005) Pressure effect studies on spin crossover systems. Coord Chem Rev 249:1811–1829
146. Woodall CH, Beavers CM, Christensen J, Hatcher LE, Intissar M, Parlett A, Teat SJ, Reber C, Raithby PR (2013) Hingeless negative linear compression in the mechanochromic gold complex [(C6F5Au)2(μ-1,4-diisocyanobenzene)]. Angew Chem Int Ed 52:9691–9694
147. Cairns AB, Goodwin AL (2015) Negative linear compressibility. Phys Chem Chem Phys 17:20449–20465
148. Madsen SR, Overgaard J, Stalke D, Iversen BB (2015) High-pressure single crystal X-ray diffraction study of the linear metal chain compound Co3(dpa)4Br2·CH2Cl2. Dalt Trans 44:9038–9043
149. Wu L-C, Jones C, Stasch A, Platts JA, Overgaard J (2014) Non-nuclear attractor in a molecular compound under external pressure. Eur J Inorg Chem 2014:5536–5540
150. Arnold PL, Prescimone A, Farnaby JH, Mansell SM, Parsons S, Kaltsoyannis N (2015) Characterizing pressure-induced uranium C−H agostic bonds. Angew Chem Int Ed 54:6735–6739
151. Jarzembska KN, Kamiński R, Dziubek KF, Citroni M, Paliwoda D, Durka K, Fanetti S, Bini R (2018) Impact of high pressure on metallophilic interactions and its consequences for spectroscopic properties of a model tetranuclear Silver(I)–Copper(I) complex in the solid state. Inorg Chem 57:8509–8520
152. Woodall CH, Christensen J, Skelton JM et al (2016) Observation of a re-entrant phase transition in the molecular complex tris(μ 2-3,5-diisopropyl-1,2,4-triazolato-κ 2 N 1:N 2) trigold(I) under high pressure. IUCrJ 3:367–376
153. Byrne PJ, Richardson PJ, Chang J, Kusmartseva AF, Allan DR, Jones AC, Kamenev KV, Tasker PA, Parsons S (2012) Piezochromism in Nickel salicylaldoximato complexes: tuning crystal-field splitting with high pressure. Chem Eur J 18:7738–7748
154. Bruce-Smith IF, Zakharov BA, Stare J, Boldyreva EV, Pulham CR (2014) Structural properties of nickel dimethylglyoxime at high pressure: single-crystal X-ray diffraction and DFT studies. J Phys Chem C 118:24705–24713
155. Takeda K, Hayashi J, Shirotani I, Fukuda H, Yakushi K (2006) Structural, optical, and electrical properties of one-dimensional bis(dimethylglyoximato)nickel(II), $Ni(dmg)_2$ at high pressure. Mol Cryst Liq Cryst 460:131–144
156. Wu L-C, Overgaard J, Madsen SR, Schmøkel MS, Brummerstedt Iversen B (2013) High-pressure single-crystal X-ray diffraction study of the photomagnetic switching complex – [Y(DMF)4(H2O)3(μ-CN)Fe(CN)5]·H2O. J Chin Chem Soc 60:929–934
157. Casati N, Macchi P, Sironi A (2005) Staggered to eclipsed conformational rearrangement of [Co2(CO)6(PPh3)2] in the solid state: an X-ray diffraction study at high pressure and low temperature. Angew Chem Int Ed 44:7736–7739
158. Parois P, Moggach SA, Lennie AR, Warren JE, Brechin EK, Murrie M, Parsons S (2010) The effect of pressure on the crystal structure of Gd(PhCOO)(3)(DMF) (n) to 3.7 GPa and the transition to a second phase at 5.0 GPa. Dalt Trans 39:7004–7011
159. Wittlinger J, Werner S, Schulz H (1998) Pressure-induced order-disorder phase transition of spinel single crystals. Acta Crystallogr Sect B 54:714–721
160. Yan H, Yang F, Pan D et al (2018) Sterically controlled mechanochemistry under hydrostatic pressure. Nature 554:505

161. Clegg JK, Brock AJ, Jolliffe KA, Lindoy LF, Parsons S, Tasker PA, White FJ (2017) Reversible pressure-controlled depolymerization of a copper(II)-containing coordination polymer. Chem Eur J 23:12480–12483
162. Prescimone A, Morien C, Allan D, Schlueter JA, Tozer SW, Manson JL, Parsons S, Brechin EK, Hill S (2012) Pressure-driven orbital reorientations and coordination-sphere reconstructions in [CuF2(H2O)2(pyz)]. Angew Chem Int Ed 51:7490–7494
163. Willett RD, Gatteschi D, Kahn O (1985) Magneto-structural correlations in exchange coupled systems. D Reidel Publishing, Boston
164. Parois P, Moggach SA, Sanchez-Benitez J, Kamenev KV, Lennie AR, Warren JE, Brechin EK, Parsons S, Murrie M (2010) Pressure-induced Jahn–Teller switching in a Mn12 nanomagnet. Chem Commun 46:1881–1883
165. Ghannadzadeh S, Möller JS, Goddard PA et al (2013) Evolution of magnetic interactions in a pressure-induced Jahn-Teller driven magnetic dimensionality switch. Phys Rev B 87:241102
166. Ovcharenko V, Romanenko G, Polushkin A et al (2019) Pressure-controlled migration of paramagnetic centers in a heterospin crystal. Inorg Chem 58:9187–9194
167. Dobe C, Strässle T, Juranyi F, Tregenna-Piggott PLW (2006) Pressure-induced switch of the direction of the unique Jahn−Teller axis of the Chromium(II) hexaqua cation in the deuterated ammonium chromium Tutton salt. Inorg Chem 45:5066–5072
168. Augustyniak MA, Krupski M (1999) The temperature dependence of the pressure switching of Jahn–Teller deformation in the deuterated ammonium copper Tutton salt. Chem Phys Lett 311:126–130
169. Prescimone A, Milios CJ, Sanchez-Benitez J et al (2009) High pressure induced spin changes and magneto-structural correlations in hexametallic SMMs. Dalt Trans 25:4858–4867
170. Craig GA, Sarkar A, Woodall CH et al (2018) Probing the origin of the giant magnetic anisotropy in trigonal bipyramidal Ni(ii) under high pressure. Chem Sci 9:1551–1559
171. Shepherd HJ, Rosa P, Vendier L, Casati N, Létard J-F, Bousseksou A, Guionneau P, Molnár G (2012) High-pressure spin-crossover in a dinuclear Fe(ii) complex. Phys Chem Chem Phys 14:5265–5271
172. Shepherd HJ, Bonnet S, Guionneau P, Bedoui S, Garbarino G, Nicolazzi W, Bousseksou A, Molnár G (2011) Pressure-induced two-step spin transition with structural symmetry breaking: X-ray diffraction, magnetic, and Raman studies. Phys Rev B 84:144107
173. Laisney J, Shepherd HJ, Rechignat L, Molnár G, Rivière E, Boillot ML (2018) Pressure-induced switching properties of the iron(iii) spin-transition complex [FeIII(3-OMeSalEen)2] PF6. Phys Chem Chem Phys 20:15951–15959
174. Shepherd HJ, Palamarciuc T, Rosa P, Guionneau P, Molnár G, Létard J-F, Bousseksou A (2012) Antagonism between extreme negative linear compression and spin crossover in [Fe(dpp)2(NCS)2]·py. Angew Chem Int Ed 51:3910–3914
175. Batten Stuart R, Champness Neil R, Chen X-M, Garcia-Martinez J, Kitagawa S, Öhrström L, O'Keeffe M, Paik Suh M, Reedijk J (2013) Terminology of metal–organic frameworks and coordination polymers (IUPAC recommendations 2013). Pure Appl Chem 85:1715
176. Bennett TD, Cheetham AK (2014) Amorphous metal–organic frameworks. Acc Chem Res 47:1555–1562
177. Gaillac R, Pullumbi P, Beyer KA, Chapman KW, Keen DA, Bennett TD, Coudert F-X (2017) Liquid metal–organic frameworks. Nat Mater 16:1149
178. Moggach SA, Bennett TD, Cheetham AK (2009) The effect of pressure on ZIF-8: increasing pore size with pressure and the formation of a high-pressure phase at 1.47 GPa. Angew Chem Int Ed 48:7087–7089
179. Serre C, Millange F, Thouvenot C, Noguès M, Marsolier G, Louër D, Férey G (2002) Very large breathing effect in the first nanoporous Chromium(III)-based solids: MIL-53 or CrIII(OH)·{O2C−C6H4−CO2}·{HO2C−C6H4−CO2H}x·H2Oy. J Am Chem Soc 124:13519–13526

180. Zhang C, Gee JA, Sholl DS, Lively RP (2014) Crystal-size-dependent structural transitions in nanoporous crystals: adsorption-induced transitions in ZIF-8. J Phys Chem C 118:20727–20733
181. Hobday CL, Woodall CH, Lennox MJ, Frost M, Kamenev K, Düren T, Morrison CA, Moggach SA (2018) Understanding the adsorption process in ZIF-8 using high pressure crystallography and computational modelling. Nat Commun 9:1429
182. Chapman KW, Halder GJ, Chupas PJ (2008) Guest-dependent high pressure phenomena in a nanoporous metal−organic framework material. J Am Chem Soc 130:10524–10526
183. Graham AJ, Tan J-C, Allan DR, Moggach SA (2012) The effect of pressure on Cu-btc: framework compression vs. guest inclusion. Chem Commun 48:1535–1537
184. Graham AJ, Allan DR, Muszkiewicz A, Morrison CA, Moggach SA (2011) The effect of high pressure on MOF-5: guest-induced modification of pore size and content at high pressure. Angew Chem Int Ed 50:11138–11141
185. Graham AJ, Banu A-M, Düren T, Greenaway A, McKellar SC, Mowat JPS, Ward K, Wright PA, Moggach SA (2014) Stabilization of scandium terephthalate MOFs against reversible amorphization and structural phase transition by guest uptake at extreme pressure. J Am Chem Soc 136:8606–8613
186. Aguado S, Bergeret G, Titus MP, Moizan V, Nieto-Draghi C, Bats N, Farrusseng D (2011) Guest-induced gate-opening of a zeolite imidazolate framework. New J Chem 35:546–550
187. Yot PG, Ma Q, Haines J et al (2012) Large breathing of the MOF MIL-47(VIV) under mechanical pressure: a joint experimental–modelling exploration. Chem Sci 3:1100–1104
188. Cavka JH, Jakobsen S, Olsbye U, Guillou N, Lamberti C, Bordiga S, Lillerud KP (2008) A new zirconium inorganic building brick forming metal organic frameworks with exceptional stability. J Am Chem Soc 130:13850–13851
189. Su Z, Miao Y-R, Zhang G, Miller JT, Suslick KS (2017) Bond breakage under pressure in a metal organic framework. Chem Sci 8:8004–8011
190. Hobday CL, Marshall RJ, Murphie CF et al (2016) A computational and experimental approach linking disorder, high-pressure behavior, and mechanical properties in UiO frameworks. Angew Chem Int Ed 55:2401–2405
191. Rogge SMJ, Wieme J, Vanduyfhuys L, Vandenbrande S, Maurin G, Verstraelen T, Waroquier M, Van Speybroeck V (2016) Thermodynamic insight in the high-pressure behavior of UiO-66: effect of linker defects and linker expansion. Chem Mater 28:5721–5732
192. Yot PG, Yang K, Ragon F, Dmitriev V, Devic T, Horcajada P, Serre C, Maurin G (2016) Exploration of the mechanical behavior of metal organic frameworks UiO-66(Zr) and MIL-125(Ti) and their NH2 functionalized versions. Dalt Trans 45:4283–4288
193. Bennett TD, Tan J-C, Moggach SA, Galvelis R, Mellot-Draznieks C, Reisner BA, Thirumurugan A, Allan DR, Cheetham AK (2010) Mechanical properties of dense zeolitic imidazolate frameworks (ZIFs): a high-pressure X-ray diffraction, nanoindentation and computational study of the zinc framework Zn(Im)2, and its lithium boron analogue, LiB(Im)4. Chem Eur J 16:10684–10690
194. Dissegna S, Vervoorts P, Hobday CL, Düren T, Daisenberger D, Smith AJ, Fischer RA, Kieslich G (2018) Tuning the mechanical response of metal–organic frameworks by defect engineering. J Am Chem Soc 140:11581–11584
195. Wu H, Yildirim T, Zhou W (2013) Exceptional mechanical stability of highly porous zirconium metal–organic framework UiO-66 and its important implications. J Phys Chem Lett 4:925–930
196. Peterson GW, DeCoste JB, Glover TG, Huang Y, Jasuja H, Walton KS (2013) Effects of pelletization pressure on the physical and chemical properties of the metal–organic frameworks Cu3(BTC)2 and UiO-66. Microporous Mesoporous Mater 179:48–53
197. Peterson GW, DeCoste JB, Fatollahi-Fard F, Britt DK (2014) Engineering UiO-66-NH2 for toxic gas removal. Ind Eng Chem Res 53:701–707

Struct Bond (2020) 185: 199–238
https://doi.org/10.1007/430_2020_78

Published online: 11 October 2020

Watching Photochemistry Happen: Recent Developments in Dynamic Single-Crystal X-Ray Diffraction Studies

Lauren E. Hatcher, Mark R. Warren, Anuradha R. Pallipurath, Lucy K. Saunders, and Jonathan M. Skelton

Contents

L. E. Hatcher
School of Chemistry, Cardiff University, Cardiff, UK
e-mail: HatcherL1@cardiff.ac.uk

M. R. Warren and L. K. Saunders
Diamond Light Source Ltd., Diamond House, Harwell Science and Innovation Campus, Didcot, UK

A. R. Pallipurath
School of Chemical and Process Engineering, University of Leeds, Leeds, UK

J. M. Skelton (✉)
Department of Chemistry, University of Manchester, Manchester, UK
e-mail: jonathan.skelton@manchester.ac.uk

Abstract Photoresponsive materials are an important contemporary research area with applications in, for example, energy and catalysis. Mechanistic information on solid-state photochemical reactions has traditionally come from spectroscopy and modelling, with crystallography limited to snapshots of endpoints and long-lived intermediates. Recent advances in X-ray sources and detectors have made it possible to follow solid-state reactions in situ with dynamic single-crystal X-ray diffraction (SCXRD) methods, allowing a full set of atomic positions to be determined over the course of the reaction. These experiments provide valuable structural information that can be used to interpret spectroscopic measurements and to inform materials design and optimisation.

Solid-state linkage isomers, where small-molecule ligands such as NO, NO_2^-, N_2 and SO_2 show photo-induced changes in binding to a transition metal centre, have played a leading role in the development of dynamic SCXRD methodology, since the movement of whole atoms and the predictable temperature dependence of the excited-state lifetimes make them ideal test systems. The field of "photocrystallography", pioneered by Coppens in the late 1990s, has developed alongside advances in instrumentation and computing and can now provide the 3D structures of species with lifetimes down to femtoseconds.

In this chapter, we will review the development of photocrystallography experiments against linkage isomer systems, from the early identification of metastable species under continuous illumination, through measuring kinetics at low temperature, to recent experiments studying species with sub-second lifetimes. We will discuss the advances in X-ray sources and instrumentation that have made dynamic SCXRD experiments possible, and we will highlight the role of kinetic modelling and complementary spectroscopy in designing experiments. Finally, we will discuss possible directions for future development and identify some of the outstanding challenges that remain to be addressed.

Keywords Lasers · LEDs · Linkage isomers · Metastable states · Photochemistry · Photocrystallography · Reaction kinetics · Solid state · X-ray diffraction

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
bipy	2,2′-Bipyridine
BPh_4	Tetraphenylborate
Bu_4dien	*N,N,N′,N′*-Tetrabutyldiethylenetriamine
CCD	Charge-coupled device

CMOS	Complementary metal-oxide-semiconductor
dcpe	1,2-*Bis*(dicyclohexylphosphino)ethane
dppe	1,2-*Bis*(diphenylphosphino)ethane
ES	Excited state
Et_4dien	*N,N,N′,N′*-Tetraethyldiethylenetriamine
GS	Ground state
JMAK	Johnson-Mehl-Avrami-Kolmogorov
LED	Light-emitting diode
MS_1	Metastable state 1
MS_2	Metastable state 2
MX	Macromolecular crystallography
OTf	Trifluoromethanesulfonate
SCXRD	Single-crystal X-ray diffraction
SFX	Serial femtosecond crystallography
SNP	Sodium nitroprusside
TR	Time resolved
XFEL	X-ray free-electron laser

1 Introduction

Modern time-resolved (TR) spectroscopy with pulsed lasers can routinely access nanosecond and femtosecond timescales. However, unless spectroscopic signatures can be assigned to transient species, for example, using theoretical modelling, interpreting the data from these studies can be challenging. Recent advances in X-ray sources and detectors make possible an alternative: time-resolved single-crystal X-ray diffraction (TR-SCXRD) [1]. A TR-SCXRD experiment yields a sequence of snapshots of the full 3D structure of the crystal over the course of a solid-state reaction. This allows for transient species or intermediates to be identified to a high level of confidence without a priori knowledge, provided they are present at more than a few % of the crystal volume. State-of-the-art serial crystallography measurements pioneered by the macromolecular crystallography (MX) community using X-ray free-electron laser (XFEL) sources can access femtosecond dynamics, while experiments using synchrotron and laboratory X-ray light sources can be used to study species with lifetimes from picoseconds to seconds and longer, allowing a wide range of light-activated solid-state reactions to be visualised in 3D.

While TR-SCXRD methods were pioneered by the MX community [2], the earliest in situ dynamic SCXRD studies on molecular crystals were performed in the 1990s to investigate the unexpectedly long-lived metastable excited states in transition metal linkage isomers [1], which have since served as model systems for more recent developments.

Linkage isomerism is a phenomenon where a ligand can display two or more distinct binding modes to a transition metal centre. The first reports of linkage

isomerism date back to the early 1900s, when the existence of yellow and red forms of the coordination complex $[Co(NH_3)_5(NO_2)]Cl_2$ was explained by the nitrite (NO_2^-) ligand coordinating through N or O to produce two complexes with distinct crystal field splitting (Fig. 1). In 1956, it was shown using infrared (IR) spectroscopy that the O-bound nitrito form converts to the more stable N-bound nitro form over time in the solid state [3]. Solution ^{18}O labelling experiments further demonstrated that the NO_2 ligand does not exchange with solvent and thus that the isomerisation is an intramolecular process [4].

A breakthrough in this field came in 1997 when Carducci et al. reported SCXRD experiments on sodium nitroprusside ($Na_2[Fe(CN)_5(NO)].2H_2O$; SNP), in which single crystals were photoexcited in situ using a laser [1]. Previous studies using Mössbauer spectroscopy and calorimetry showed that light irradiation at low temperature produced two long-lived metastable intermediates that can co-exist in the same crystal. The crystallography experiments confirmed the existence of three distinct isomers corresponding to different binding modes of the NO ligand, viz. a

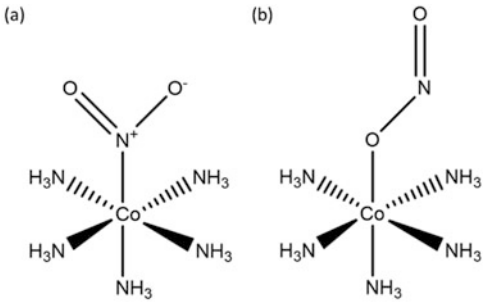


Fig. 1 Linkage isomers of the $[Co(NH_3)_5(NO_2)]^{2+}$ cation. The yellow N-bound nitro isomer (**a**) is the most energetically stable, but the red O-bound nitrito isomer (**b**) can be obtained as a kinetic product

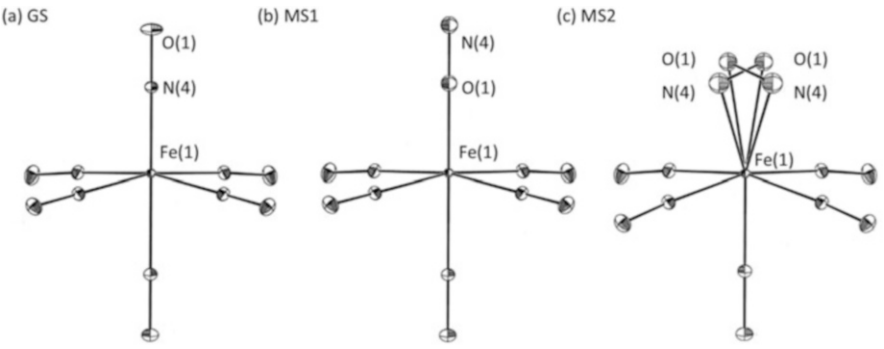


Fig. 2 X-ray structures of the ground-state (GS) and major and metastable state (MS1/MS2) isomers of the nitroprusside anion $[Fe(CN)_5(NO)]^{2-}$. Adapted with permission from Ref. [1]. Copyright 1997 American Chemical Society

ground-state NO isomer, an excited-state ON isomer and a second side-bound excited-state isomer (Fig. 2), with the different metastable isomers accessed by varying the irradiation wavelength. The combination of SCXRD and in situ photoexcitation was coined "photocrystallography" and is now commonly used to describe this research area.

In this chapter, we will review the historical development and current state of the art in photocrystallography in the context of linkage isomers. We will start from early photocrystallographic experiments to identify metastable species and proceed to discuss the development of low-temperature dynamic SCXRD experiments to probe the isomerisation kinetics. We will then finish with a discussion of how modern synchrotron facilities and photon-counting detectors allow the structures of species with shorter lifetimes to be determined.

This chapter is organised as follows. In Sect. 2 we will give a brief overview of the major classes of linkage isomer systems and some of the key research studies in each area, and in Sect. 3, we will introduce the photocrystallography experiments used to study them. Finally, in Sect. 4 we will discuss the main considerations for designing time-resolved SCXRD experiments to follow the isomerisation processes in real time, leading to recent developments towards sub-second SCXRD experiments on linkage isomers.

2 Linkage Isomer Systems

Several families of linkage isomer systems have been identified with ligands including NO_2^- (nitrite), NO (nitrosyl), sulphoxide (SO_2) and (di-)nitrogen (N_2). Some of the known binding configurations of these ligands in transition metal complexes are summarised in Fig. 3.

In most of these systems, the complexes crystallise with the ligand in its energetic ground-state geometry, and can be excited into one or more metastable configurations by photoactivation. Metastable isomers typically have long lifetimes at cryogenic temperatures and are thus kinetically trapped, but decay rapidly back to the ground state on warming to room temperature.

2.1 Nitrite (NO_2^-) Systems

Following the discovery of $[Co(NH_3)_5(NO_2)]^{2+}$ in the 1850s [5], nitrites remain one of the best studied classes of linkage isomer materials to date. Though a number of coordination geometries are theoretically possible for the nitrite ligand [6], in practice the monodentate nitro (η^1-NO_2), *endo*-nitrito (η^1-ONO) and *exo*-nitrito (η^1-ONO) arrangements are the most commonly observed in molecular crystals. Photochemical conversion between the nitro and nitrito isomers of $[Co(NH_3)_5(NO_2)]^{2+}$ in solution was first reported in the 1940s [7], while solid-state measurements were first performed on microcrystalline powders in the 1970s. As

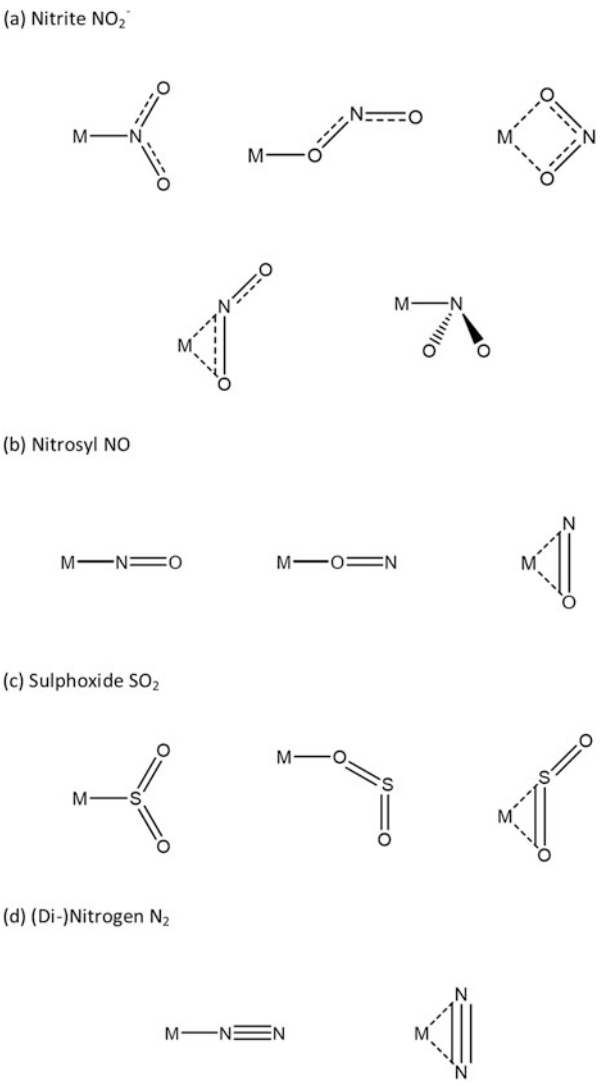


Fig. 3 Binding modes of common photo-switchable ligands found in solid-state linkage isomer systems to a metal centre (M): (**a**) nitrite (NO_2^-), (**b**) nitrosyl (NO), (**c**) sulphoxide (SO_2) and (**d**) dinitrogen (N_2)

photocrystallographic studies became established in the late 1990s and early 2000s, transition metal nitrite systems have played a key role in furthering our understanding of linkage isomerism in the single crystal.

Much of the work on solid-state nitrite linkage isomers to date has aimed to maximise the nitro → nitrito photoconversion level by engineering a "reaction cavity" around the switchable ligand to allow the isomerisation to proceed with minimal steric influence from the surrounding crystal environment. This strategy follows from the topochemical postulate proposed by Schmidt and Cohen in the 1960s, which states that single-crystal reactions must follow the minimum-energy pathway involving the least amount of atomic rearrangement in order to preserve

crystallinity [8, 9] and which was extended to by Cohen and later Ohashi et al. [10, 11] to propose the introduction of reaction cavities to facilitate solid-state reactions. Maximising the size of the reaction cavity around the switchable ligand allows the structure to accommodate the changes during isomerisation with minimal change in the overall structure, thereby minimising crystal strain and facilitating higher conversion.

The first linkage isomer system to show 100% conversion to the metastable state in the single crystal, designed using this principle, was the Ni(II)-nitrite complex [Ni(dppe)(η^1-$\underline{N}O_2$)Cl] (dppe = 1,2-*bis*(diphenylphosphino)ethane) [12], which shows complete nitro → nitrito conversion within 90 min when irradiated with 400 nm LED light. The sterically demanding and photo-inert dppe ligand dominates the crystal packing and maximises the reaction cavity around the nitrite groups, such that photoactivation results in a small 0.77% volume expansion.

The same principle was subsequently applied to design other neutral complexes with 100% conversion, including [Ni(dppe)(η^1-$\underline{N}O_2$)$_2$], [Ni(dcpe)(η^1-$\underline{N}O_2$)$_2$] (dcpe = 1,2-*bis*(dicyclohexylphosphine)ethane) [13] and [Ni(Et_4dien)(η^2-$\underline{O},\underline{O}N$)($\eta^1$-$\underline{N}O_2$)] ($Et_4$dien = *N,N,N′,N′*-tetraethyldiethylenetriamine) [14]. Multicomponent crystals with photoactive cations and counter-anions are also known, and the possibility of incorporating sterically demanding counterions expands the design space. For example, the [Pd(Bu_4dien)(NO_2)]BPh_4 complex (Bu_4dien = *N,N,N′,N′*-tetrabutyldiethylenetriamine, BPh_4 = tetraphenylborate) achieves 100% nitro → nitrito photoconversion in under 15 min when illuminated with low-power LEDs, with just 0.14% change in the unit-cell parameters, and stands out as an example of a relatively fast single-crystal-to-single-crystal photoreaction [15] (Fig. 4).

While the "reaction cavity" approach has proved extremely successful in achieving high photoconversion in the solid state, comparison of isostructural materials has highlighted additional influences from electronic and kinetic factors.

The [Pd(Et_4dien)(NO_2)]OTf and [Pt(Et_4dien)(NO_2)]OTf (OTf = trifluoromethanesulfonate) complexes adopt the same crystal structure but show markedly different behaviour on irradiation at low temperature [16]. Irradiation at 100 K produces a mixture of the metastable *endo*-nitrito (η^1-$\underline{O}NO$) and *exo*-nitrito (η^1-$\underline{O}NO$) isomers, with different conversion levels in the two systems, whereas irradiation at higher temperatures leads to 95 and 93% conversion to the *endo*-nitrito (η^1-$\underline{O}NO$) isomer in the Pd(II) and Pt(II) systems, respectively. These differences can be rationalised by thermal expansion at higher temperatures both enlarging the reaction cavity and weakening hydrogen bonding interactions to the nitro (η^1-$\underline{N}O_2$) group in the ground-state structure. The importance of intermolecular interactions involving the isomerising ligand has also been highlighted in other systems [17–19].

Within the series of isostructural [M(Et_4dien)(NO_2)]OTf systems, isomerisation of the M = Pd(II) completes in 1 h but is dramatically slowed to 3 h in the M = Pt (II) material using the same excitation source. This appears to be a common feature of nitrite linkage isomer systems based on third-row transition metals [20] and may be rationalised in terms of the decrease in kinetic lability of the ligands in transition metal complexes when descending a group [21].

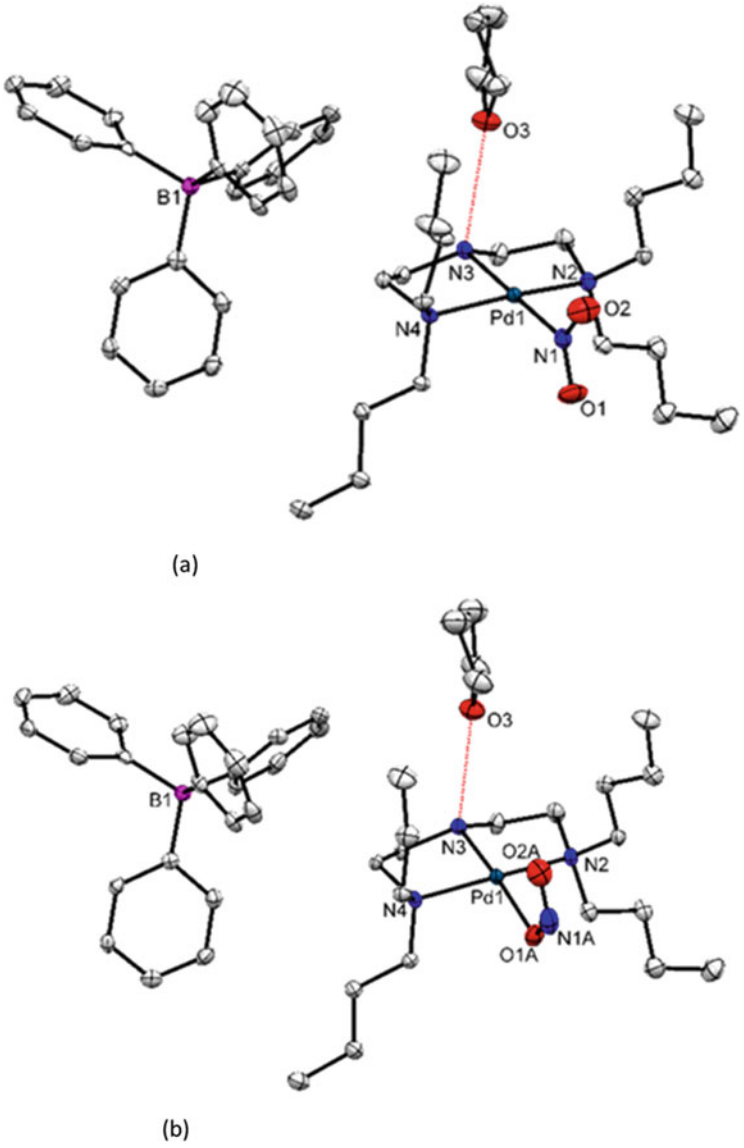


Fig. 4 Ground- and excited-state structures of $[Pd(Bu_4dien)(NO_2)]BPh_4$ at 100 K. (**a**) Structure of the ground-state nitro (η^1-NO_2) isomer. (**b**) Structure of the 100% converted excited-state *endo*-nitrito (η^1-ONO) isomer obtained after irradiation at 400 nm for 1 h. Adapted from Ref. [15] under the Creative Commons Attribution license

Finally, the importance of the electronic structure and the conditions necessary for photochemical linkage isomerism have been thoroughly investigated in metal-nitrosyl systems [22], and much of this insight is thought to be transferable to other small ambidentate ligands such as nitrite and sulphur dioxide.

2.2 Nitrosyl (NO) Systems

The seminal study that helped establish the field of photocrystallography was the work on the iron-nitrosyl complex sodium nitroprusside (SNP; Fig. 2) [1]. The 1997 study by Coppens et al. confirmed the presence of two unexpectedly long-lived metastable states in SNP, previously identified by Mössbauer spectroscopy [23], and assigned them as photogenerated linkage isomers. Irradiation of a single crystal at 50 K with 488 nm light produces a 37% conversion from the ground-state η_1-NO isomer to the inverted η_1-ON isomer (MS1). Further irradiating the partially converted crystal at 1064 nm further induces conversion to 10% of a second side-bound η_2-NO isomer (MS2).

SNP has since been used as a model system to explore the electronic landscape of the photochemical linkage isomerisation in the solid state, using a variety of structural, theoretical, spectroscopic and thermal analysis methods to explore the relationships between the three isomers [22, 24–26]. It is now believed that irradiation with visible light causes initial conversion to the side-bound MS2 form, which then converts to the O-bound MS1 isomer under prolonged exposure. The subsequent excitation at 1,064 nm then causes depopulation of MS1, which decays via MS2.

Other reported transition metal-nitrosyl linkage isomers investigated using photocrystallography include systems based on Rh [27], Ni [28] and Ru. Early work by Coppens in 1996 investigated the photogenerated MS1 and MS2 isomers in $[Ru(NO_2)_4(OH)(NO)]$ [29], which was later followed by studies on ruthenium porphyrin nitrosyls [30]. In parallel, Schaniel and Woike demonstrated photoactivation in *trans*-$[RuCl(py)_4(NO)][PF_6]_2.0.5H_2O$ using infrared (IR) spectroscopy [31], and subsequent photocrystallographic studies confirmed a 92% conversion to MS1 at 80 K on irradiation with 673 nm laser light, followed by 48% conversion to MS2 at 980 nm [32]. These successes have led to a series of related complexes with different halides, pyridine ligands and counterions, enabling the influence of the surrounding crystalline environment on the isomerisation to be established [19].

Another interesting example is the $[Ru(NO_2)(bipy)_2(NO)]$ (bipy = 2,2′-bipyridine) complex, for which photoactivation of single crystals facilitates a double isonitrosyl-nitrito linkage isomerisation involving the intramolecular transfer of an oxygen atom between the NO_2 and NO ligands [33]. This highlights the potential for competing photo-release reactions, although in practice these are more likely to be observed in solution [34] and/or in studies conducted at, or close to, room temperature. NO release has been investigated in solid-state nitrosyl materials [35, 36], but is not typically observed in nitrite and sulphur dioxide systems.

2.3 Sulphur Dioxide (SO_2) Systems

Sulphur dioxide (SO_2) linkage isomerism has been demonstrated both in solution and in the single crystal. The ground-state structure is typically an S-bound

arrangement, while irradiation induces conversion to O-bound η^1-OSO MS1 or side-bound η^2-(O,SO) MS2 isomers similar to the nitrosyl linkage isomer species. The first evidence for sulphur dioxide linkage isomerism was obtained in 1979 from solution spectroscopy and proposed two distinct metastable isomers [37]. The side-bound isomer was confirmed in 2002 from steady-state photocrystallographic measurements on *trans*-$[Ru(NH_3)_4Cl(SO_2)]Cl$ with 300–500 nm light at temperatures below 150 K [38, 39]. The MS1 isomer was confirmed by Bowes et al. in 2006 from photocrystallographic measurements on $[Ru(NH_3)_4(H_2O)(SO_2)][MeC_6H_4SO_3]_2$, which shows 36% conversion to the η^1-OSO form on irradiation at 13 K [40].

From these early studies, several other ruthenium sulphur dioxide compounds have been designed and studied using photocrystallography to establish the effects of the crystal environment on the isomerisation process. In particular, systematic studies on $[Ru(SO_2)(NH_3)_4X]Y$, varying both the ligand X and the counter-anion Y, have shown that the excited-state geometry is strongly influenced by the crystal environment and the nature of the *trans* ligand [41–43]. More recent studies have also demonstrated photoisomerisation in osmium SO_2 complexes [44].

2.4 Dinitrogen (N_2) Systems

Though comparatively less studied, photoinduced linkage isomerism can also occur with dinitrogen (N_2) ligands. N_2 binds to transition metal centres in both mono- and bidentate coordination geometries (c.f. Fig. 3), and the earliest reports of dinitrogen coordination complexes date back to the 1960s [45]. A variety of complexes with terminal and bridging N_2 ligands have since been investigated, with recent interest in their use for catalytic nitrogen fixation [46–48].

Single-crystal dinitrogen linkage isomerism was first demonstrated in the osmium complex $[Os(NH_3)_5(N_2)](PF_6)_2]$ [49], for which photocrystallographic measurements showed that a maximum of 17.4% of the molecules in the crystal could be converted from a ground-state "end-on" η^1-N_2 arrangement to a side-bound η^2-N,N isomer by irradiation with 325 nm laser light at 100 K. The side-bound isomer was confirmed to be metastable and found to persist for at least a day at this temperature.

2.5 Linkage Isomer Device Development

The design of linkage isomers with high conversion has been motivated by the possibility of using them in molecular devices. For example, one could consider a molecular data storage application similar to an optical disc where the state of a binary digit (bit) is encoded by the majority isomer (ground state or excited state; GS/ES) in small spots on a thin film. Bits are set by photoactivating a spot with a laser, reset by localised heating and read with a suitable probe. To be able to write

data quickly, we require rapid and (near-)complete GS → ES conversion on illumination, and we would ideally want the conversion to be achievable with low power. To maintain data integrity, the excited state should be metastable – perhaps indefinitely – at a specified operating temperature, which would ideally be at or near room temperature to avoid the need for elaborate cooling systems. A third property that may be desirable is the presence of a "reporter" that could be used to read out the majority isomer in a region of the recording surface. This might, for example, be a discriminatory infrared (IR) absorption, a change in the UV or visible absorption at a well-defined wavelength or a quantifiable change in a bulk physical property (e.g. the refractive index).

Though in practice linkage isomer materials are still a considerable way from being implemented in real devices, several studies have successfully demonstrated how this could be achieved. In 2010, Woike et al. demonstrated that the change in the refractive index due to photoisomerisation in three different $[Fe(CN)_5(NO)]^{2-}$ systems could be used to write phase gratings in the crystals, showcasing their potential use for holographic data storage [50]. In a different approach, Cole et al. have investigated the use of ruthenium sulphur dioxide linkage isomers as molecular motors and demonstrated photo-induced nanomechanical transduction in crystals of $[Ru(SO_2)(NH_3)_4(3\text{-chloropyridine})]Y_2$ (Y = *p*-tosylate or 4-chlorobenzenesulfonate) [51]. Here, photoactivation leads to conversion to both the MS1 and MS2 isomers, depending on the irradiation wavelength, and an additional rotation of the benzene ring in the counter-anion is observed alongside the MS1 conversion as a result of changes in intermolecular interactions. Incorporation of ruthenium sulphur dioxide materials into thin films of polyvinyl alcohol (PVA) has also been demonstrated [52], showing how single crystals might be encapsulated into device media.

3 Steady-State and Pseudo-Steady-State Photocrystallographic Methodology

In this section we will illustrate how a typical photocrystallography experiment is conducted in the laboratory. We will also introduce the Johnson-Mehl-Avrami-Kolmogorov (JMAK) kinetic model for the photoactivated excitation and thermal decay processes, which will form the basis for discussing faster pump-probe experiments in Sect. 4.

3.1 Experimental Setup

In a typical photocrystallography experiment, a crystal is mounted on the X-ray diffractometer and irradiated in situ using a light source. In a laboratory setting, integrating a high-power pulsed laser setup is often impractical, and so lower-power

irradiation sources such as low-power continuous-wave lasers or LEDs are common. Of these, LED setups are an attractive choice as they allow for flexible array design, a choice of a wide variety of LED wavelengths and form factors and are comparatively low cost.

An example of an LED setup is shown in Fig. 5 [53]. In this setup the LEDs are mounted in a supporting ring attached to the end of the diffractometer cryostream. The ring positions six LEDs with the centre of the emission cone directed to the sample position and incorporates gaps for the incident and diffracted X-ray beams. Other LED device designs for in situ SCXRD incorporate fibre optics and focussing mirrors to achieve the desired illumination arrangement [54]. As well as being highly practical and easy to implement, LED arrays also spread the radiation load uniformly over time and provide better spatial coverage across the crystal surface, avoiding potential problems with laser beam damage, unwanted sample heating and uneven illumination.

3.2 Static Ground and Excited States (Photostationary Structures)

Where the excited state has a lifetime of hours or greater at low temperature, standard SCXRD experiments can be conducted to establish the ground- and excited-state structures.

A crystal is mounted and cooled in the dark, allowing a structure of the ground state to be obtained with minimal contamination from the excited state. The crystal is then illuminated for a set amount of time, which typically ranges from minutes to hours depending on the system. At low temperatures, the excited state is kinetically trapped, and a substantial steady-state population accumulates during irradiation. This typically happens at temperatures below 150 K, although the critical temperature varies between systems. A second X-ray dataset is then collected to obtain the structure of the excited state. To confirm that a maximum excitation level has been reached, a second period of illumination is usually performed and inspected for further conversion to the excited-state isomer. Once the maximum photostationary population has been obtained, the crystal is held in the dark, and a series of variable-temperature data collections are completed to investigate the temperature range over which the metastable state decays.

Figure 4 in the previous section shows as an example of the structures of the ground-state nitro (η^1-$\underline{N}O_2$) and 100% excited-state *endo*-nitrito (η^1-$\underline{O}NO$) isomers of the $[Pd(Bu_4dien)(NO_2)]BPh_4$ nitrite linkage isomer complex.

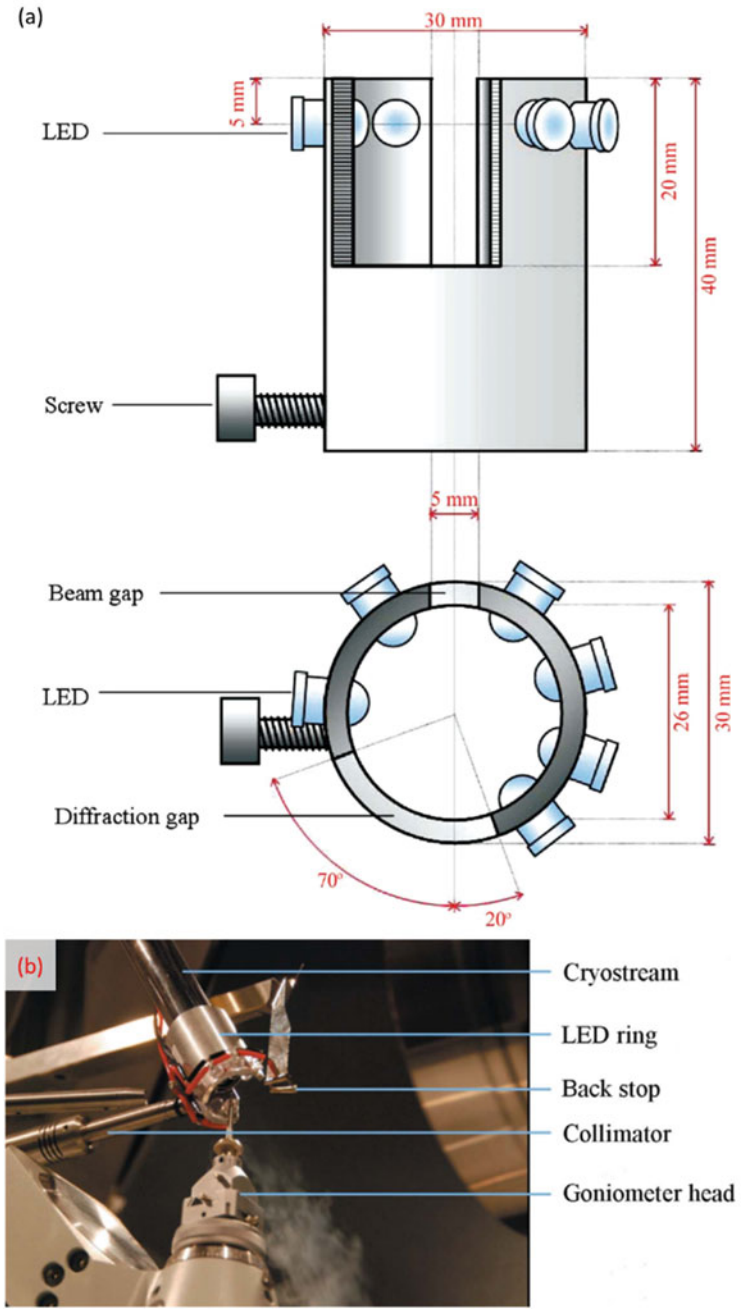


Fig. 5 Example of an LED setup for in situ irradiation of crystals on a laboratory X-ray diffractometer. (**a**) Technical drawing of an LED holder with six LEDs illuminating the sample and gaps for the incident and diffracted X-ray beams. (**b**) Photograph of the LED holder mounted at the sample position by attachment to the cryostream. Reproduced from Ref. [53] with permission from the International Union of Crystallography

3.3 Excitation Kinetics Measurements

The kinetics of the excitation process (i.e. the excited-state population dynamics) in the single crystal can be measured by irradiating the sample at low temperature and collecting an X-ray structure after each exposure. Provided the excited-state population shows no appreciable decay at low temperature over the timescale of a data collection, it is possible to measure the growth of the excited-state population with irradiation time in this manner. Figure 6 shows a schematic illustration of a typical excitation experiment.

Figure 7 shows an excitation measurement performed on a $[Pd(Bu_4dien)(NO_2)]BPh_4$ crystal [15]. In this experiment, a single crystal was irradiated with 400 nm LED light at 100 K, and structures were collected after 1, 2, 5, 10 and 15 min of total exposure.

The evolution of the excited-state population with exposure time shows a characteristic exponential behaviour that can be well described by the Johnson-Mehl-Avrami-Kolmogorov (JMAK) model [55–57]. The JMAK model predicts the

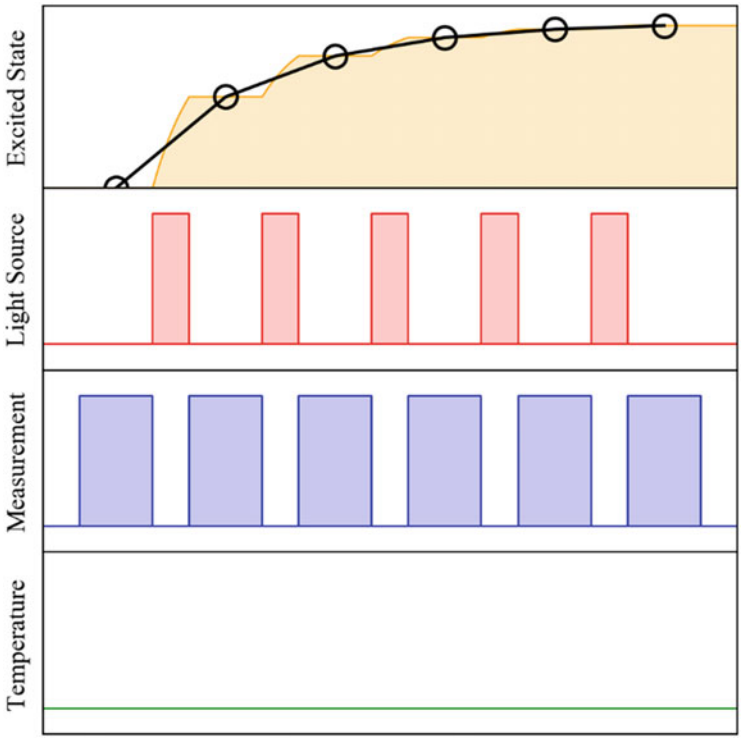


Fig. 6 Schematic illustration of a typical excitation measurement. The experiment is performed at low temperature where the thermal decay of the excited state is negligible. Starting from a clean ground state, excitation pulses and X-ray measurements are interleaved to record the change in the excited-state population as a function of total irradiation time

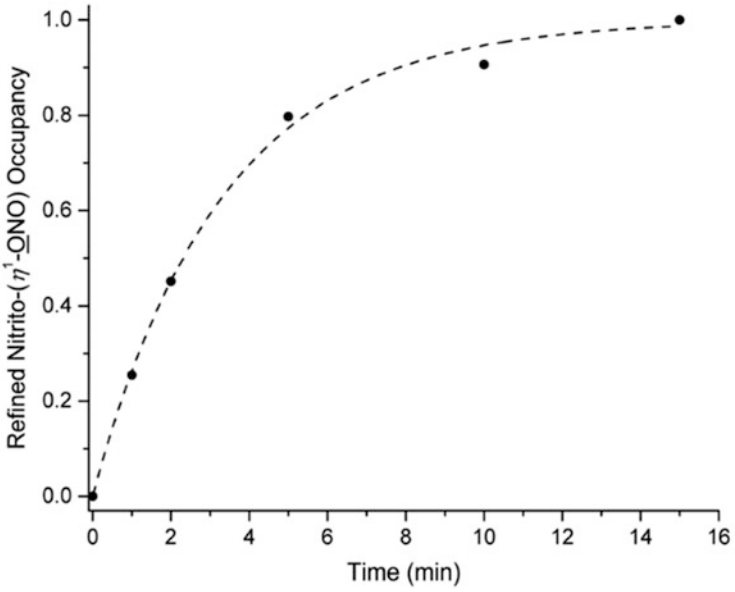


Fig. 7 Evolution of the fractional population of the excited-state *endo*-nitrito (η^1-ONO) isomer of the [Pd(Bu_4dien)(NO_2)]BPh_4 linkage isomer system at 100 K as a function of irradiation time. Reproduced from Ref. [15] under the Creative Commons Attribution license

volume fraction $\alpha(t)$ of the excited-state isomer at a time t from the start of the excitation as:

$$\alpha(t) = 1 - \exp(-kt^n) \tag{1}$$

where k is a rate constant and n is the (usually integer) Avrami exponent. n is related to the dimensionality D of the transformation and provides information about the level of cooperativity:

$$D = n - 1 \tag{2}$$

$n = 1$ ($D = 0$) implies non-cooperative, homogenous excitation throughout the bulk of the crystal. For materials such as [Pd(Bu_4dien)(NO_2)]BPh_4 that have been designed according to the "reaction cavity" concept outlined in Sect. 2, the isomerisation takes place in a cavity shielded by bulky ligands and induces minimal strain in the surrounding crystal bulk. The excitation is therefore expected to be non-cooperative, and the results of JMAK analysis with these materials tend to confirm this by predicting a value of n~1 and therefore simple exponential behaviour [15, 58, 59].

A value of $n > 1$ would indicate cooperative transformation in 1 ($n = 2$), 2 ($n = 3$) or 3 ($n = 4$) spatial dimensions. Such behaviour has been found for single-crystal-to-

single-crystal [2 + 2] cycloaddition reactions, where dimerisation of one molecule pre-organises neighbouring molecules for better orbital overlap and more efficient dimerisation, lowering the energetic barrier to conversion and thereby propagating a 1D chain reaction [60, 61].

It is worth noting that the excitation rate k depends on the power and wavelength of the excitation source and will also depend on the size and shape (morphology) of the crystal and its physical properties including absorption cross-section, refractive index and any light-scattering and/or reflection effects. The electronic absorptions that lead to the photoisomerisation are typically charge-transfer bands close to the onset of the UV-vis absorption, and the band positions and band widths determine whether the photoexcitation can be induced over a narrow or wide distribution of wavelengths.

For efficient photoexcitation it may seem optimal to select a wavelength close to the absorption maximum where the absorption cross-section is largest. However, this needs to be balanced against the penetration depth of the light into the crystal bulk given by the Beer-Lambert law:

$$I(r) = I_0 \exp(-\tau r) \tag{3}$$

where $I(r)$ is the intensity at a depth r from the surface in the crystal bulk, I_0 is the incident intensity and τ is the attenuation. The intensity falls exponentially with the distance from the surface, and the attenuation τ is set by the absorption cross-section. Efficient excitation and penetration depth are thus competing factors, and an optimum balance is usually obtained by selecting wavelengths close to the tail of the absorption band [62, 63].

For completeness, a third factor that may need to be taken into account is that if the absorption spectrum of the excited-state red shifts relative to the ground state, an outer layer of converted molecules may prevent the light from reaching unconverted molecules at the core and thus limit the reaction rate (and therefore k) and the maximum conversion level. On the other hand, the exponent n should be an intrinsic property of the material, at least under the assumption that different excitation wavelengths or powers do not access fundamentally different excitation mechanisms.

3.4 Decay Kinetics Measurements

Once suitable conditions for excitation have been established, measurements can be made to investigate the decay kinetics. As shown in Fig. 8, a typical decay experiment occurs in two steps. First, the crystal is cooled to low temperature and irradiated to build up a large (ideally 100%) initial excited-state population. In the second step, the excitation source is switched off, the sample is warmed rapidly to the target measurement temperature, and the decay of the excited-state

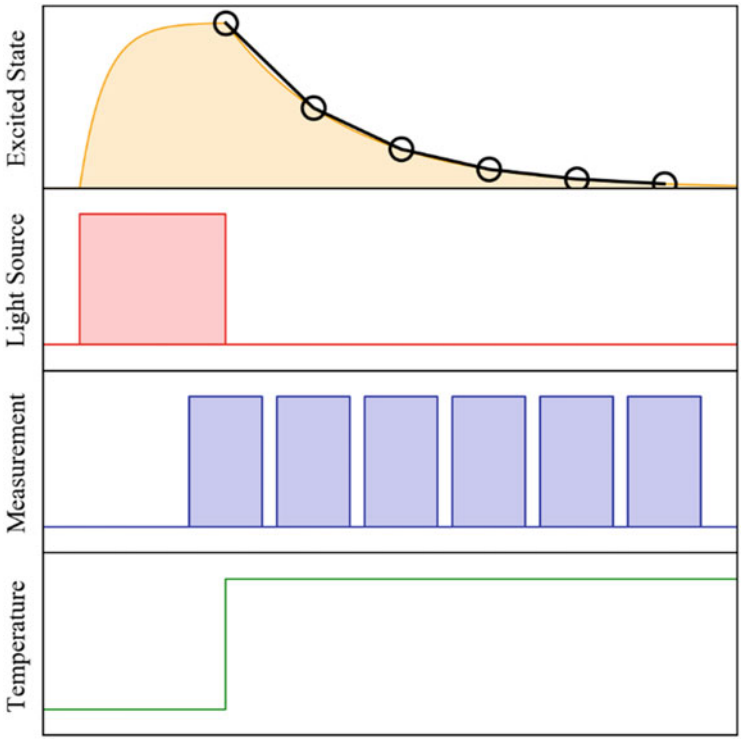


Fig. 8 Schematic of a typical decay measurement. The crystal is first irradiated to generate a large initial excited-state population. The temperature is then raised, and the change in population over time is monitored by performing continuous X-ray measurements until the thermal decay is complete

population with time is monitored by collecting a continuous series of X-ray datasets.

Figure 9a shows a series of decay measurements on a single crystal of the [Pd(Bu_4dien)(NO_2)]BPh_4 system. As in the excitation process, the exponential decay behaviour is well modelled using a different form of the JMAK equation:

$$\alpha(t) = \exp(-kt^n) \tag{4}$$

where k is again the rate constant for the decay process and n indicates the level of cooperativity and is typically close to 1 (non-cooperative decay) for the same reasons as the excitation.

The measurements in Fig. 9a show that the thermal decay of the nitrito-(η^1-ONO) isomer of [Pd(Bu_4dien)(NO_2)]$^+$ is very strongly temperature dependent, such that the initial 100% excited-state population decays to zero within 180 min (3 h) at 230 K, but is only ~60% complete after 720 min (12 h) at 212.5 K. The temperature-

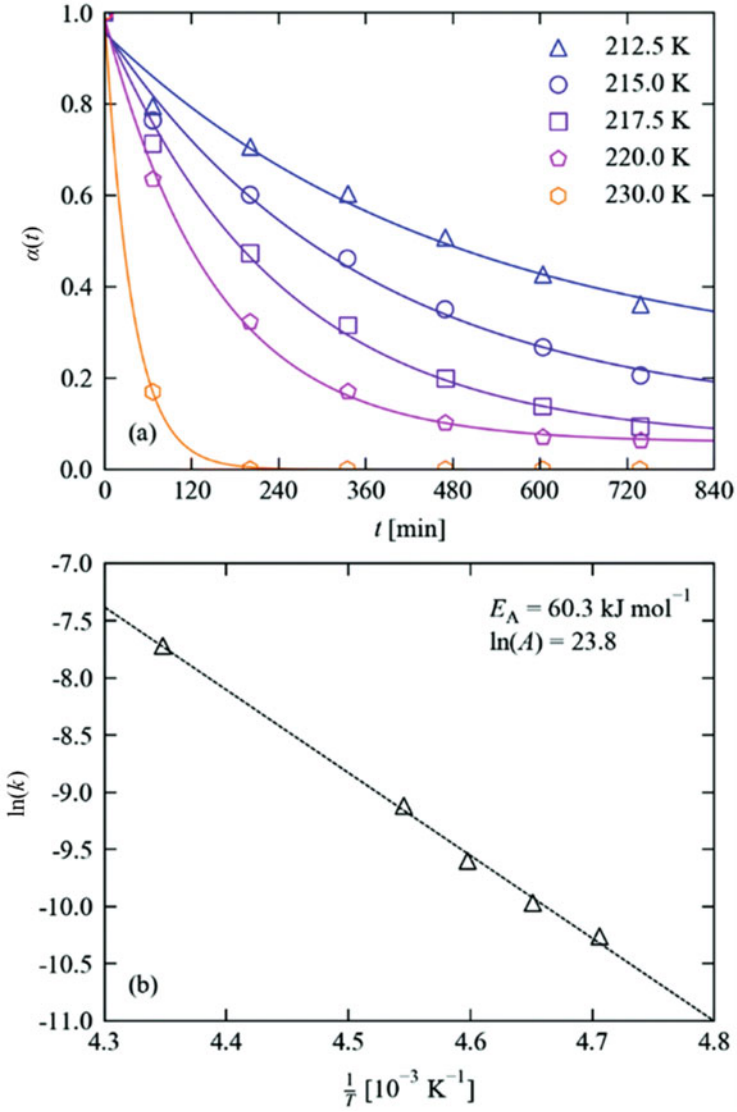


Fig. 9 (**a**) Decay of the fractional population of the excited-state η^1-ONO (nitrito) isomer of the [Pd(Bu_4dien)(NO_2)]BPh_4 linkage isomer system at temperatures between 212.5 and 230 K. The markers show data points obtained by refining single-crystal X-ray structures recorded during the decay, and the solid lines show fits to the JMAK expression in Eq. (4) with $n = 1$. (**b**) Temperature dependence of the fitted rate constants. The markers show the data points, and the dashed line shows a fit to the linearised Arrhenius law in Eq. (6) with the parameters $E_A = 60.3$ kJ mol^{-1} and $\ln(A) = 23.8$. Adapted from Ref. [59] under the Creative Commons Attribution license

dependent decay rate is a common feature of linkage isomer systems and is usually well modelled using the Arrhenius law:

$$k(T) = A\ \exp\left(-\frac{E_A}{RT}\right) \tag{5}$$

where E_A is the activation energy associated with the isomerisation and the pre-exponential factor A has the same units as k. A can be loosely interpreted as an attempt frequency and the exponential term as the probability that each attempt leads to successful isomerisation. It is worth noting that for a typical linkage isomer system with $n = 1$, the rate constant has units of s^{-1} and A therefore has frequency units.

If a series of decay measurements are performed at different temperatures and fitted to the JMAK model to extract a set of rate constants, the linearised form of the Arrhenius form can then be used to obtain an estimate for the attempt frequency and activation energy for the decay process:

$$\ln\left[k(T)\right] = \ln A - \frac{E_A}{R}\frac{1}{T} \tag{6}$$

If $k(T)$ follows the Arrhenius law, $\ln[k(T)]$ and the inverse temperature $1/T$ show a linear relationship and E_A and A can be determined from the slope and intercept, respectively.

The Arrhenius plot in Fig. 9b shows that the $k(T)$ extracted from the decay data in Fig. 9a can be fitted to the Arrhenius law with a large activation energy of 60.3 kJ mol^{-1} and $\ln(A) = 23.8$, which corresponds to an attempt frequency of 2.17×10^{10} Hz = 21.7 GHz. The attempt frequency is too low to be associated with, e.g. a metal-ligand bond stretch, which would be on the order of tens of THz, but may be a low-energy lattice vibration or a ligand rotation. However, this should be treated with caution, since $\ln(A)$ is obtained from the intercept by extrapolation.

It is interesting to examine the temperature dependence of the decay process by relating the activation energy to the excited-state lifetime. We begin with the more general form of the JMAK equation:

$$\alpha(t) = \alpha_\infty + (\alpha_0 - \alpha_\infty)\exp\left[-kt^n\right] \tag{7}$$

α_0 and α_∞ correspond to the excited-state populations at $t = 0$ and $t = \infty$, respectively [64]. Substituting $\alpha_0 = 0$ and $\alpha_\infty = 1$ into Eq. (7) recovers the excitation expression in Eq. (1) for the ideal situation of complete excitation, while substituting $\alpha_0 = 1$ and $\alpha_\infty = 0$ recovers the expression for complete decay in Eq. (4). This more general form of the JMAK equation is useful for modelling systems where either the excitation or decay process never reaches completion.

Assuming that $\alpha_0 = 1$ and $\alpha_\infty = 0$, we can obtain an expression for the time $t_{1/2}$ required for the initial excited-state population to fall to $\alpha = 0.5$ by rearranging Eq. (7):

$$\exp\left[-k\left(t_{1/2}\right)^{n}\right] = 0.5 \rightarrow t_{1/2} = \left[-\frac{1}{k}\ln 0.5\right]^{\frac{1}{n}} \tag{8}$$

In the case of $n = 1$, $t_{1/2}$ can be equated to the kinetic half-life, meaning that a population α at a time t will fall to $\alpha/2$ at $t = t + t_{1/2}$, $\alpha/4$ at $t = t + 2t_{1/2}$, etc. $t_{1/2}$ thus provides a useful quantitative measure of the excited-state lifetime. Its temperature dependence can be expressed in terms of the Arrhenius parameters A and E_A by substituting for the rate constant k as follows:

$$t_{1/2}(T) = \left[-\ln 0.5 \times \frac{1}{A} \times \exp\left(\frac{E_A}{RT}\right)\right]^{\frac{1}{n}} \tag{9}$$

This expression makes three important points concerning the excited-state lifetime: (1) the lifetime increases with the activation energy; (2) the lifetime decreases with the attempt frequency; and (3) the lifetime decreases as the inverse power of the JMAK exponent n. The first two points are straightforwardly understood from the Arrhenius model. The third point may not be as immediately intuitive, but indicates that any cooperativity in the decay process should be expected to decrease the excited-state lifetime.

Figure 10 shows the temperature dependence of $t_{1/2}$ evaluated using Eq. (9), with the Arrhenius parameters $E_A = 60.3$ kJ mol^{-1} and $\ln(A) = 23.8$ obtained for the

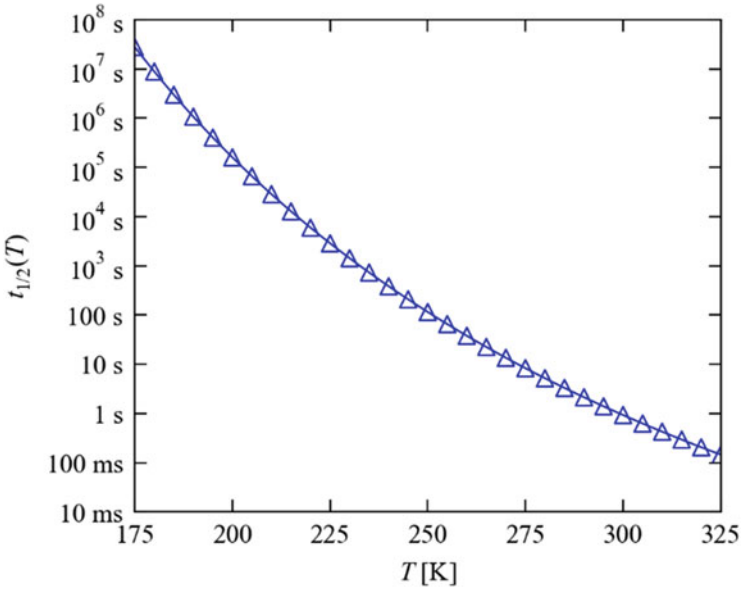


Fig. 10 Temperature dependence of the half-life $t_{1/2}$ of the excited-state η^1-ONO (nitrito) isomer of the [Pd(Bu_4dien)(NO_2)]BPh_4 linkage isomer system calculated from Eq. (9) based on the Arrhenius parameters in Fig. 9b and a JMAK exponent $n = 1$

$[Pd(Bu_4dien)(NO_2)]BPh_4$ system from the fit in Fig. 9b [59]. Rather strikingly, the excited-state lifetime varies by nine orders of magnitude, from ~10^8 s at 175 K to 100 ms at 325 K. This range of timescales may be compared to the time required to collect a full X-ray dataset. For a well-diffracting crystal, a typical laboratory X-ray system with a microfocus or rotating anode source and a CCD detector might take in the region of 1–2 h to collect a full dataset and could therefore reasonably access decay measurements up to ~230 K as in Fig. 10. A high-flux synchrotron source, on the other hand, may take only a few minutes to collect a dataset, extending the range of temperatures that can be measured up to around 250 K. Towards room temperature, the lifetime is on the order of seconds, which would require pump-probe measurements to access as we discuss in the next section.

3.5 Pseudo-Steady-State Measurements

To perform the excitation and decay measurements outlined in the previous subsections, the experimental conditions (i.e. temperature and illumination) are adjusted so as to study the excitation and decay processes independently. If the crystal is illuminated continuously at a temperature where the two processes are competitive, a steady-state equilibrium will be reached. This experiment is illustrated schematically in Fig. 11.

Slightly different behaviour is obtained depending on whether a continuous or pulsed excitation source is used. In the experiment depicted in Fig. 11a, the sample is illuminated continuously, and the excited-state population builds smoothly to a steady-state value determined by the balance of the excitation and decay rates at the measurement temperature. In Fig. 11b, on the other hand, a pulsed illumination source is used, and the excited-state population builds during the light pulses and decays between them, oscillating about an equilibrium value. Provided the repetition rate of the excitation pulses is much shorter than the timescale of the structural measurements, the refined structure will show the averaged value. In the photocrystallographic literature, the term "pseudo-steady-state" is usually used to describe both of these scenarios, although when discussing the kinetics, it is useful to distinguish the behaviour obtained with pulsed and continuous excitation sources as, e.g. "pseudo-steady-state" and "steady-state", respectively.

In most systems, the decay rate increases with temperature, while the excitation rate remains approximately constant, with the result that the steady-state excited-state population falls from a maximum to zero over a range of temperatures. Figure 12 shows a set of pseudo-steady-state measurements on the $[Pd(Bu_4dien)(NO_2)]BPh_4$ system. In this experiment, the decay becomes competitive with the excitation at around 230 K. Between 230 and 290 K, the increasing decay rate causes the steady-state excited-state population to drop from 100% to zero, and above 290 K the excited-state population is no longer measurable.

This behaviour can be understood intuitively as follows. For a given excitation wavelength and power, the excitation rate is approximately constant, or at best is

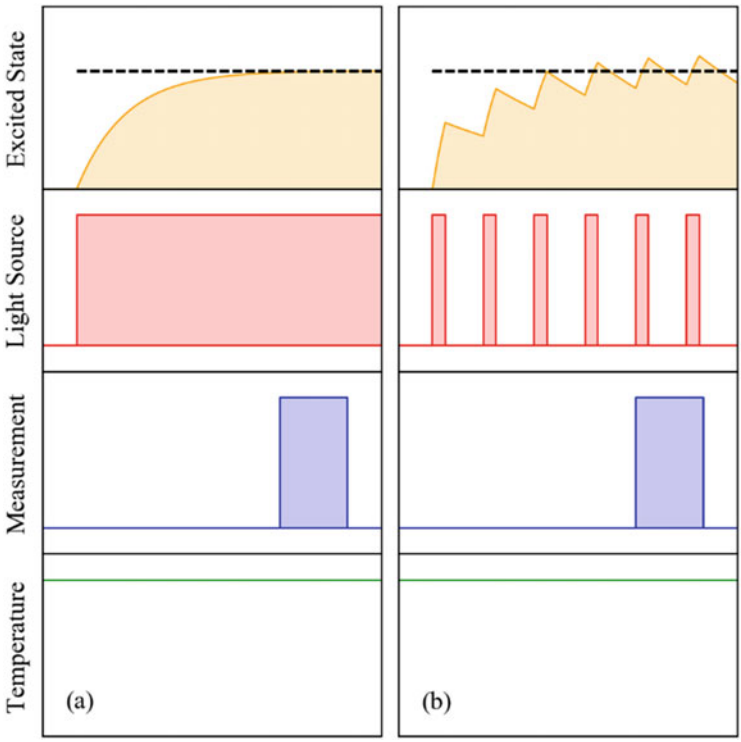


Fig. 11 Schematic of typical steady-state (**a**) and pseudo-steady-state (**b**) experiments. In (**a**), the sample is illuminated continuously, and the excited-state population builds smoothly to a steady-state value. This equilibrium population, marked by a dashed black line, is then measured after a predetermined equilibration time. In (**b**), the sample is illuminated with a pulsed excitation source which results in the excited-state population oscillating about an equilibrium value. If the measurement time is slower than the repetition rate, the average population, again denoted by the dashed black line, is measured

weakly temperature dependent [59]. The decay rate, on the other hand, is strongly temperature dependent, as per the Arrhenius law. At low temperature, the excitation is much faster than the decay, so the steady-state excited-state population approaches its maximum. As the temperature is raised, the decay becomes faster and begins to compete with the excitation, causing the steady-state population to drop. At high temperatures, the decay is much faster than the excitation, any photo-induced isomerisation is almost instantaneously reversed, and the steady-state excited-state population falls to zero.

Unlike $t_{1/2}$, it is not possible to derive a simple closed expression for the steady-state occupation as a function of temperature from JMAK and Arrhenius models. However, this can be done relatively straightforwardly using numerical simulation.

We first make the assumption that over a short time interval, Δt, the decay and excitation act independently. Given JMAK parameters for the excitation and decay

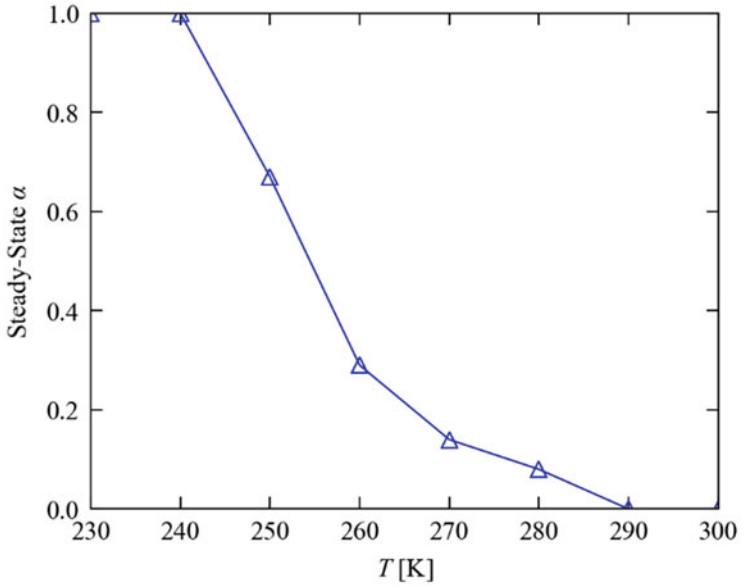


Fig. 12 Steady-state population of the excited-state *endo*-nitrito (η^1-ONO) isomer of the [Pd(Bu_4dien)(NO_2)]BPh_4 linkage isomer system under continuous illumination at 400 nm as a function of temperature. This data was taken from Ref. [59]

and a current metastable-state population $\alpha(t)$, Eqs. (1 and 4) can be rearranged to obtain the effective time t' capturing the progress of the two processes:

$$\alpha(t) = 1 - \exp\left(-k_{\text{exc}} t'_{\text{exc}}\right) \Rightarrow t'_{\text{exc}} = -\frac{1}{k_{\text{exc}}} \ln\left[1 - \alpha(t)\right] \tag{10}$$

$$\alpha(t) = \exp\left(-k_{\text{dec}} t'_{\text{dec}}\right) \Rightarrow t'_{\text{dec}} = -\frac{1}{k_{\text{dec}}} \ln \alpha(t) \tag{11}$$

Here, we have distinguished the k and t' for the two processes with subscript dec and exc and assumed $n = 1$ for simplicity.

Having obtained t'_{exc} and t'_{dec}, we can then predict the change in excited-state population $\Delta\alpha$ over the time interval $t \rightarrow t + \Delta t$ as follows:

$$\begin{aligned}\Delta\alpha(t \rightarrow t + \Delta t) = &\left\{1 - \exp\left[-k_{\text{exc}}\left(t'_{\text{exc}} + \Delta t\right)\right] - \alpha(t)\right\} \\ &- \left\{\alpha(t) - \exp\left[-k_{\text{dec}}\left(t'_{\text{dec}} + \Delta t\right)\right]\right\}\end{aligned} \tag{12}$$

The first term in braces calculates the growth in excited-state population over the interval due to excitation, while the second term calculates the reduction in population due to decay. An adjustment is then applied to clamp $\alpha(t)$ to the range [0, 1]. Provided a sufficiently small Δt is chosen, these simulations provide a good description of the time evolution of the excited-state population for a given set of conditions.

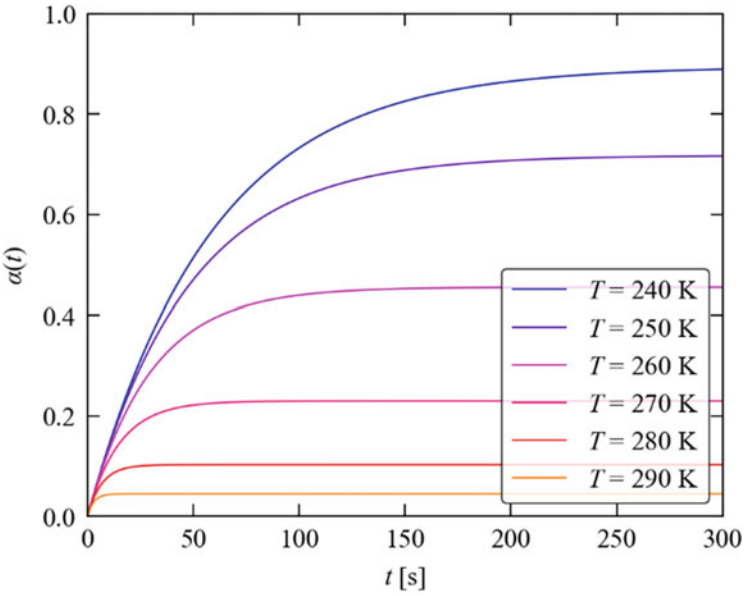


Fig. 13 Numerical simulations of the equilibration of the steady-state excited-state population α under continuous illumination at temperatures from 240 to 290 K. The simulations are parameterised with an excitation rate of $k_{exc} = 1.53 \times 10^{-2}$ s^{-1} and a temperature-dependent decay rate defined by the Arrhenius parameters $E_A = 60.3$ kJ mol^{-1} and $\ln(A) = 23.8$, based on the data from Ref. [59]

Figure 13 shows simulations of steady-state measurements where an initial excited-state population $\alpha = 0$ is equilibrated under a fixed $k_{exc} = 1.53 \times 10^{-2}$ s^{-1} at a range of temperatures where the k_{dec} is set using the Arrhenius parameters $E_A = 60.3$ kJ mol^{-1} and $\ln(A) = 23.8$ (Ref. [59], Fig. 9). As well as predicting the expected fall in the steady-state population with temperature, these simulations also show that the smaller equilibrium excited-state population is reached in a much shorter time, decreasing from around 5 min at 240 K to a few seconds below ~270 K. This is due to the exponential nature of both processes resulting in rapid excitation and slow decay at small α. The steady-state population and the time taken to reach it are both important parameters for designing successful time-resolved experiments.

Finally, it is worth noting that in some cases, pseudo-steady-state experiments can identify transient species in addition to the ground-state and excited-state isomers [58]. Like the [Pd(Bu$_4$dien)(NO$_2$)]BPh$_4$ system, the photoisomerisable nitrite ligand in the [Ni(Et$_4$dien)(η^2-O,ON)(η^1-NO$_2$)] complex excites from a nitro (η^1-NO$_2$) to an *endo*-nitrito (η^1-ONO) isomer under illumination (Fig. 14). The steady-state occupation of the *endo*-nitrito isomer drops to zero from ~140 to 190 K, but under the steady-state conditions, a second *exo*-nitrito (η^1-ONO) isomer is observed as a minor species at 5–10% occupation. Molecular dynamics simulations have shown that the *exo*- and *endo*-isomers can interconvert rapidly due to thermal motion [65].

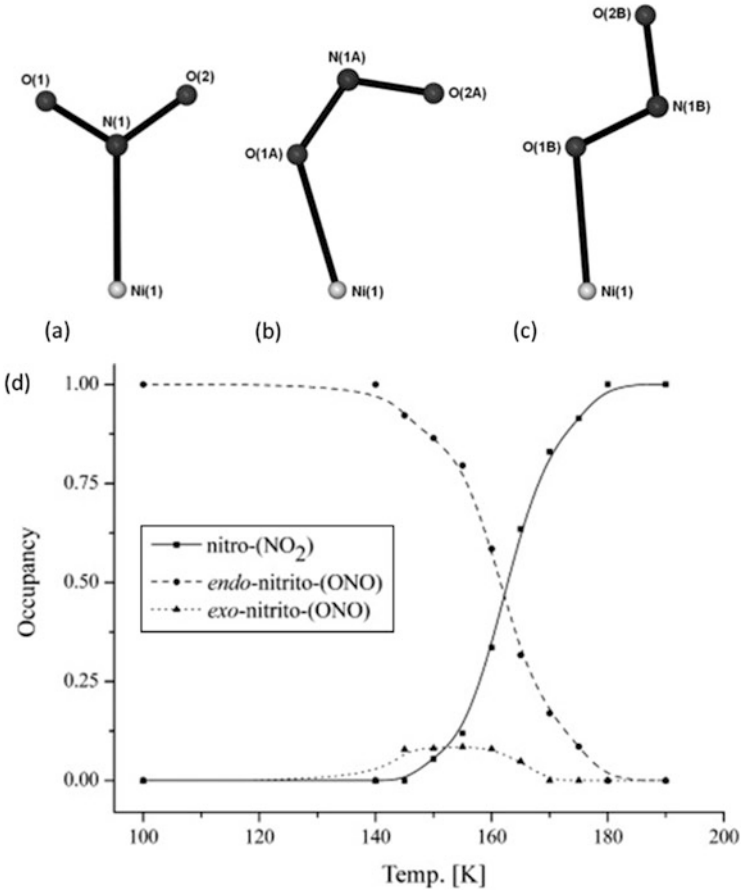


Fig. 14 Steady-state measurements on the [Ni(Et$_4$dien)(η^2-O,ON)(η^1-NO$_2$)] linkage isomer system. The photoisomerisable NO$_2$ ligand can adopt the nitro (η^1-NO$_2$) (**a**) and *endo*- and *exo*-nitrito (η^1-ONO) (**b**, **c**) binding modes. The steady-state occupation of the excited-state *endo*-nitrito isomer falls from 100% to zero from ~140 to 180 K, and the transient *exo*-nitrito isomer is observed in levels up to ~10% over this temperature window (**d**). Adapted with permission from Ref. [58]

4 Considerations for Pump-Probe Photocrystallography

Section 2 highlighted the important role that research on linkage isomer systems has played in developing modern photocrystallographic experiments. The detailed understanding of the isomerisation kinetics enabled by these measurements, outlined in Sect. 3, further highlights their potential as model systems for developing TR crystallography experiments. Systems such as the [Pd(Bu$_4$dien)(NO$_2$)]BPh$_4$ complex can be excited rapidly to high conversion levels, and the thermal decay can be adjusted to a wide range of timescales simply by changing the measurement temperature.

The goal of a TR-SCXRD experiment is to monitor the time evolution of the excited state during pulsed excitation by collecting full X-ray structures averaged over suitably short time windows. The key benefit of conducting a TR-SCXRD experiment, in comparison to more established TR spectroscopic measurements, is its ability to provide full structural information at each time point. For monochromatic TR-SCXRD studies, this means that the 3D structures of all reactants, products and short-lived intermediate species can be determined to better than atomic-scale resolution. These TR snapshots can then be combined into 3D "molecular movies", providing a unique visualisation tool that allows us to watch the solid-state reaction occurring in real time. It should be noted, however, that only transient species with lifetimes longer than, or on the order of, the experiment time resolution will be observed. Nonetheless, the insight provided by these experiments has the potential to revolutionise our understanding of photochemical reaction mechanisms, which in turn can inform the targeted design of new systems with enhanced functionality.

In this section, we explore some of the practicalities of TR-SCXRD experiments to study molecular species with lifetimes in the microsecond to millisecond range, including the choice of excitation and X-ray sources, X-ray detectors, sample delivery and data processing, and we also highlight some recent work in this area using linkage isomer crystals. Photocrystallographic experiments to study solid-state species with sub-microsecond lifetimes are discussed in chapter "Time-Resolved Single-Crystal X-Ray Crystallography".

4.1 Pump-Probe Versus Pump-Multiprobe Measurements

TR experiments are generally designed around a pump-probe measurement sequence. In the typical experiment illustrated in Fig. 15a, a regular repeating cycle is set up during which the crystal is excited by a short light pulse (the pump) at $t = 0$ and is then allowed to return to the ground state during a decay period before the next pulse arrives. A measurement (the probe) is synchronised to the pump pulse such that data is collected after a fixed time delay Δt. The short pump pulse generates a small excited-state population, and the following probe measures the population over a short time period in line with the target time resolution.

For a TR-SCXRD experiment, the recorded diffraction intensity from the short X-ray probe is typically weak, and, as a result, the pump-probe cycle must be repeated many times to obtain sufficient signal-to-noise ratio. This process must then be repeated with the crystal in many different orientations to collect complete data for structure solution. In addition, each experiment records a single time delay, and thus multiple experiments must be conducted at different Δt to build up a picture of the behaviour across the whole cycle. Taken together, these requirements can lead to overall experiment times totalling many hours or even days, depending on the required time resolution, the diffraction power of the sample and its crystal symmetry.

With the advent of electronically gated detectors with fast readout times, it is possible to perform "pump-multiprobe" measurements, where multiple Δt are

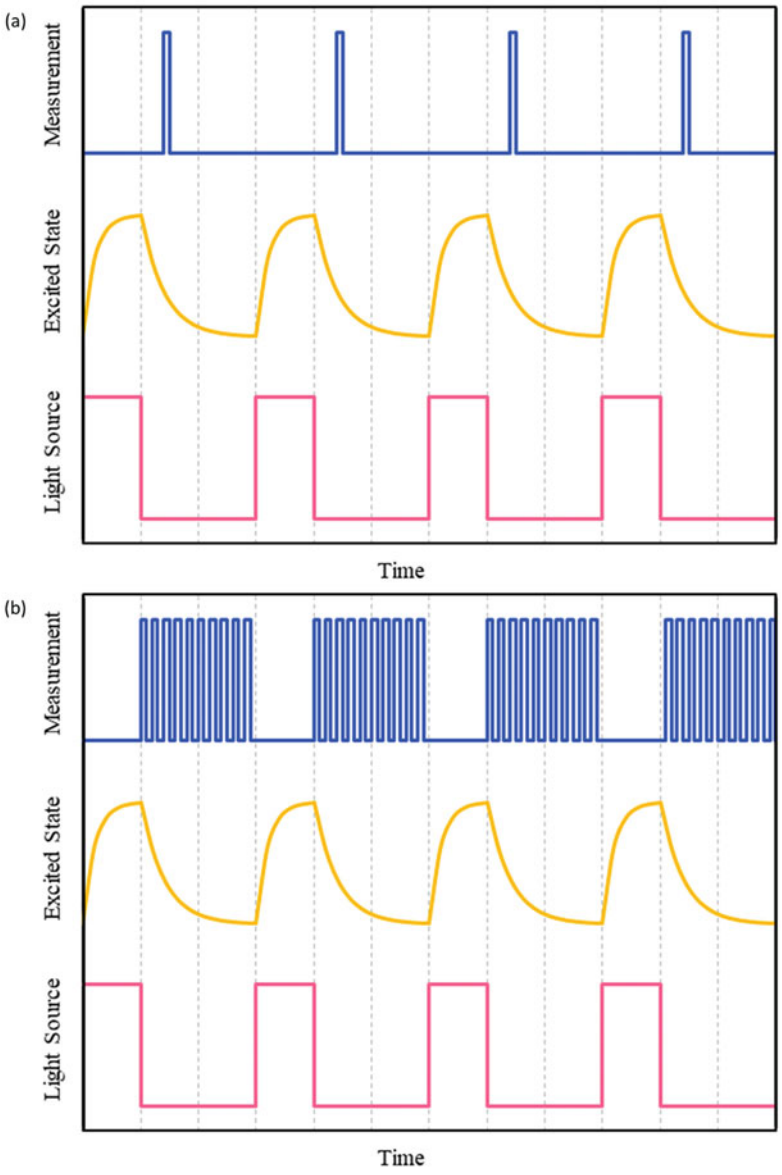


Fig. 15 Schematic of typical pump-probe and pump-multiprobe experiments. In the conventional pump-probe method shown in (**a**), a single probe pulse is timed to measure the excited-state population a specified time delay Δt after excitation with the pump light source. The whole experiment must be repeated for each Δt to be measured. In the pump-multiprobe method shown in (**b**), a series of probe pulses are generated after each pulse, allowing multiple time delays to be recorded in a single experiment. With a pulsed laser source, the pump pulse is determined by the laser pulse width and is likely to be very short. With an LED pump, the pulse width is likely to be much longer and can be varied using the controlling electronics

measured during a single experiment as shown in Fig. 15b. The process is conceptually similar to the traditional pump-probe experiment except that all the Δt are collected in one series after arrival of the excitation pulse. This removes the need for multiple pump-probe experiments to access different time delays and thus dramatically shortens the overall experiment duration.

4.2 Excitation Sources

Historically, one of the major goals of pump-probe experiments has been to achieve the shortest possible time resolution. As a result, experiments have generally used pulsed lasers to provide intense, focussed light pulses with pulse widths in the nanosecond to picosecond range. For experiments with target time resolutions in this range, lasers remain the light source of choice.

Intense laser pulses can however present difficulties that should be considered when planning an experiment. Photocrystallography experiments intentionally use the smallest crystals possible, within the limit of achieving sufficient diffraction intensity from the chosen X-ray setup, to maximise the penetration depth of the pump light through the crystal. As a consequence, a high-power laser pump with a short pulse duration will likely deliver many more photons per pulse than the number of molecular absorbers present in the sample volume. This increases the likelihood of non-linear responses such as multiphoton absorption processes that may be both unexpected and undesirable [66]. Intense light pulses may also lead to photobleaching, so that only a fraction of the photons induce photoconversion and the remaining photons are either not absorbed or, worse, produce laser heating and accelerate the decay of the excited state and/or damage the crystals.

Laser sources also bring with them several practical issues. The requirement for careful and often time-consuming alignment, together with the stringent safety protocols that need to be put in place to mitigate risks to users, adds to the complexity of the experiment. The cost of both the laser and of the optics needed to deliver the beam to the sample is generally considerable. Finally, some pulsed lasers operate at a fixed repetition rate, which somewhat reduces flexibility. If the timescale of the process being studied is much faster than the time between repetitions, considerable "dead time" is introduced between measurements, whereas if a system does not relax quickly enough between pulses, it may simply not be possible to study it with the setup. Newer, variable repetition rate lasers circumvent this issue, but are considerably more expensive.

If a high intensity and short pulse duration are not required, it may be worthwhile considering other excitation sources. As outlined in Sect. 3, some photocrystallographic experiments, particularly on laboratory equipment, are performed using LEDs. LEDs are cheap and readily available in a range of visible and near-UV wavelengths, can reach a reasonably high power output and can be

powered and driven by inexpensive electronics. They can also be flexibly built into custom arrays to fit the diffraction setup, allowing irradiation of multiple sides of the crystal. This is a distinct benefit over point laser sources, as irradiation at multiple crystal faces may enhance the achievable photoconversion level. Paired with a comparatively inexpensive function generator, LEDs can also be pulsed down to microseconds or faster. Using LEDs in this way provides fine control over the pump-probe cycle time, including the ability to vary the total cycle time and the portion of the time for which the LEDs are active. Where the target time resolution allows, lengthening the LED pump-pulse (c.f. Fig. 15) may be beneficial for some systems as this helps to spread out the energy load on the crystal in time, potentially circumventing some of the heating/damage and absorption issues outlined above and, in principle, accessing higher conversion levels.

4.3 X-Ray Sources

The X-ray source and detector primarily determine how much signal can be accumulated during each pump-probe (or pump-multiprobe) cycle. A high-flux X-ray source and efficient detector will allow the same signal-to-noise ratio to be obtained in fewer cycles and a shorter overall experiment duration than with a lower-flux source or less efficient detector. As well as the absolute time saving, a shorter experiment minimises the total exposure of the sample to the excitation source, which is important for systems where excitation causes progressive crystal damage or bleaching.

Laboratory X-ray instruments are based on the Coolidge hot-cathode X-ray tube. A tungsten filament cathode is heated to glowing temperature under vacuum to produces free electrons by thermionic emission. The electrons are accelerated at very high voltage and directed by an electrostatic lens to impact a metal anode. On impacting the anode, the majority of the electron kinetic energy is dissipated as heat, but a small percentage is converted X-ray photons. Electrons deflected by atomic nuclei produce a continuous spectrum of polychromatic X-rays through the bremsstrahlung ("braking radiation") effect. Provided the kinetic energy is greater than the binding energy, some impacts can eject core electrons to produce vacant core holes. When these are filled by relaxation of valence electrons, the excess energy is emitted as X-ray photons with well-defined energies. By filtering the emission spectrum to remove the bremsstrahlung radiation using monochromator crystals, monochromatic X-rays at these atomic wavelengths are obtained.

Typical X-ray tubes for crystallography operate at 50 kV and use Cu or Mo anodes for which the strongest K α emission lines are at 8.046 and 17.480 keV (1.54 and 0.71 Å, respectively). The improved availability of Ag has also led to sources using Ag K α radiation becoming more common in the laboratory (22.163 keV, 0.56 Å). The X-ray flux produced by the tube depends on the power

and the efficiency with which the primary electron beam produced at the cathode is converted to X-rays by the anode material. The power is a product of the cathode emission current and the accelerating voltage. The conversion efficiency is proportional to the accelerating voltage and atomic number Z of the anode material and is around 1.3, 1.9 and 2.1% for Cu, Mo and Ag tubes operating at 50 kV. The power a tube can be operated at is limited by the heat load on the anode, which must be carefully controlled to avoid rapid degradation. The anode is therefore typically water cooled.

Dramatic technical advances in this area in the last 30 years have led to substantial improvements in the brilliance of laboratory X-ray sources. Microfocus X-ray sources, capable of focussing the X-ray beam down to a much smaller spot size than conventional sources, can enhance the brilliance for smaller crystals by ~20× compared to a traditional sealed-tube source. Rotating anode sources, where the anode is moved to spread the electron beam over a larger section of the material to reduce the heat load and allow the tube to operate at higher power, can produce ~60–120× higher brilliance when combined with microfocus optics. Other new technologies such as liquid metal target and nanofocus sources offer even higher brilliance, and continuing technological advances in this area make it likely that even brighter laboratory X-ray sources will be available in the near future [67].

Despite the promising improvements in laboratory X-ray technology, as discussed in Sect. 4.1, pump-(multi)probe SCXRD measurements are severely diffraction-limited due to the need for very short X-ray probe times. At shorter target time resolutions, brighter X-ray sources are thus required to achieve sufficient signal-to-noise ratio in a reasonable timeframe. As such, most TR-SCXRD studies have been conducted at third-generation synchrotron sources, where the average X-ray brilliance is more than a billion times higher than a laboratory X-ray source [67, 68]. The advent of X-ray free electron laser (XFEL) sources has further increased the peak brilliance achievable by a further billion- to trillion-fold, enabling even smaller crystals to be studied at very short timescales.

The required time resolution also dictates the choice of X-ray source. Unlike laboratory sources, the manner in which synchrotron and XFEL radiation is generated results in a natural temporal structure [69–71]. Synchrotron sources produce X-rays in bunches with bunch widths in the region of picoseconds and typical separations between bunches of a few nanoseconds within the main bunch train, and single bunches can be isolated in the hybrid fill patterns to obtain a couple of hundred nanosecond separation to the main bunch train. This structure places a lower limit on the achievable time resolution. At nanosecond time resolution, experiments can be conducted by selecting single X-ray bunches from the synchrotron beam, while for experiments at microsecond and longer timescales, the synchrotron beam appears continuous, and techniques such as mechanical chopping or electronically gating the detector can be used to produce probes of the desired width. Further discussion of fast TR diffraction experiments at synchrotrons can be found in chapter “Time-Resolved Single-Crystal X-Ray Crystallography”. XFEL facilities provide access to even shorter timescales, with current sources producing X-ray pulses as short as a few femtoseconds and developments targeting attosecond time resolution.

4.4 X-Ray Detectors

Current X-ray detectors used for crystallography are based on one of three technologies: charge-coupled device (CCD), complementary metal-oxide-semiconductor (CMOS) or hybrid photon-counting (HPC) systems.

CCDs are based on early digital imaging technology. Early detectors used a layer of scintillator material (commonly referred to as a phosphor) to convert incident X-ray photons into visible photons, which were then directed through fibre optics to an array of photodiodes to convert the photons to electron/hole pairs. More modern detectors use materials such as silicon, CdTe and GaAs that absorb X-rays and produce electron-hole pairs directly (termed "direct detection"). The electrons are captured in potential wells and accumulated. When the image is read out, the charges in each well are transferred sequentially to an amplifier, and the signals are digitised and processed to measure the intensity of X-ray photons received at each pixel and form a monochrome intensity image.

While well established, CCDs have some notable downsides [72]. Firstly, the readout process is slow, taking on the order of a second to read out an image, and the X-rays must be shuttered while this happens. Secondly, thermally generated electrons are also captured in the charge wells, producing a high level of background noise referred to as a dark current, and mitigating this typically requires actively cooling the detector. The transfer of charges between wells to the amplifier during readout can also lead to signal loss and introduce noise. Finally, the charge wells have a finite capacity that can be exceeded during long exposure times or when measuring high-intensity reflections. This both saturates the intensity of the pixel and causes excess electrons to spill over into adjacent wells (an effect termed "blooming"), leading to distorted peak shapes. The significant background noise makes it difficult to reliably measure weak reflections, while saturation/blooming limits the dynamic range and makes it difficult to measure images that contain both very strong and very weak reflections.

CMOS detectors again evolved from digital photography and improve upon CCD detectors by attaching addressable amplification and readout logic to each pixel. By digitising the signal at source and removing the need to transfer charges between wells during readout, the signal is higher, and CMOS detectors can run at higher working temperatures with less noise than CCDs. The per-pixel logic also makes it possible to correct for issues including detector non-linearity, dark current and detector inhomogeneity at the level of individual pixels [73]. The pixels are separated with guards that prevent blooming. Although the readout time for a complete image remains similar to CCDs, addressing individual pixels allows pixels to be read out in lines to implement a so-called rolling shutter which, if images are carefully reconstructed after capturing, can considerably reduce the collection time. On the other hand, saturation remains an issue, limiting the dynamic range, and incorporating electronics into each pixel requires them to be physically larger than on CCD detectors.

HPC detectors operate on a different principle to CCD and CMOS detectors and function as arrays of independent point detectors [72, 74]. The detector surface consists of a pixelated substrate made of doped silicon, GaAs or CdTe that absorbs X-rays and produce electron-hole pairs in proportion to the X-ray photon energy. The pixels are biased to drive the electrons or holes through a metal interconnect to a readout chip. (The "hybrid" in the name refers to the way in which the sensor and readout chip are made separately from different materials and later bonded to create a working detector.) The readout chip compares the electrical signal to pre-set thresholds and increments a counter if the signal is within the threshold. Compared to CCD and CMOS detectors, X-ray photons are detected directly with no analogue-to-digital conversion, and if the thresholds are calibrated for the X-ray energy, the detector can effectively reject background noise.

While HPC detectors address the primary deficiencies of CCD and CMOS detectors, the technology nonetheless has some limitations. As with CMOS detectors, the integrated electronics require large pixels. Amplifying and detecting a signal blocks ("paralyses") the detector for a short period of time – typically hundreds of nanoseconds – during which it cannot count additional photons. Each pixel therefore has a maximum count rate beyond which it will effectively saturate, which is typically on the order of 10^5–10^7 counts per pixel per second. The operating principle also means that if two photons are captured by a pixel at the same time, only one will be counted. For completeness, it is worth noting that the digital counters in the pixels are limited to a maximum by the bit depth, but this is in practice much higher than for CCD and CMOS detectors and is only likely to be an issue for long exposures at high flux. For example, the 20-bit counter used on some PILATUS detectors can count up to $2^{20} \approx 10^6$ photons per pixel during a single exposure, which can be compared to a typical ~10^4–10^5 photons per pixel for CCD and CMOS detectors.

A second important feature of HPC detectors is the vastly improved readout speed. The counters are individually addressable by design and can be read out in parallel, which makes it possible to read out images in milliseconds, i.e. up to three orders of magnitude faster than CCD/CMOS detectors. In most detectors, the readout electronics and counter are connected through logic gates that can be electronically controlled (gated) by an external trigger signal, allowing them to be active for periods of time on the order of hundreds of ns, much shorter than the readout time.

For completeness, we note that the count rate limits of HPC detectors make them unsuitable for use with XFEL sources. As a result, several XFELs use detectors based on CCD and CMOS technologies, while others use modified HPC detectors that integrate the electronic signal over time [75, 76].

Finally, a very recent development to HPC detector technology changes the fundamental way the data is read out. Rather than counting photons to generate a 2D image, these detectors output a continuous data stream of "events" in which each detected photon is labelled with a timestamp and the 2D coordinates of the pixel that detected it [77]. This completely eliminates the readout time, and the stream may also incorporate external trigger signals for synchronisation with, e.g. a pulsed laser.

Once the experiment has been performed, the data stream can be processed to reconstruct 2D images, if needed, or, for example, to count events in a region of interest (e.g. a single diffraction peak). The key advantages of this readout method are that (1) it eliminates detector readout time and (2) it allows the time resolution of the experiment to be decided upon and adjusted after data collection, without repeating the experiment.

It can be seen from this discussion that – with the possible exception of experiments using XFEL sources – HPC detectors are much better suited to TR experiments than conventional CCD and CMOS detectors. As well as providing better quality data, including, importantly, the ability to accurately measure strong and weak reflections in a single image, HPC detectors can also provide much better time resolution, and this can be adjusted to suit different types of experiment. Most synchrotron sources now use HPC detectors, and despite their higher cost they are becoming more commonplace on laboratory instruments. The recent developments towards continuous readout are even more interesting, since these effectively push the time resolution close to the limit of the count rate and into the microseconds to hundreds of nanoseconds range.

4.5 Sample Delivery

For many TR-SCXRD experiments, obtaining a complete time series from the one crystal may not be feasible. Samples may degrade with exposure to high-intensity pump light and/or high-flux X-ray beams, and structural changes over many repeated cycles may lead to accumulation of strain and fatigue. Crystal damage is a particularly important consideration in experiments using high-brilliance X-ray sources such as synchrotrons or XFELs and high-intensity pulsed lasers, though ideally should be considered in any photocrystallographic experiment regardless of the instrumentation and intended time resolution. Also, if an irreversible reaction is being studied, then it is clearly not possible perform complete pump-(multi)probe experiments on the same crystal.

These situations call for rapid replacement of the sample and what are commonly termed “multi-crystal methods”. Multi-crystal methods have mainly been developed to cater to serial femtosecond crystallography (SFX) experiments conducted at XFEL sources but are also applicable to synchrotron studies. During XFEL experiments the high-intensity X-ray beam effectively destroys the crystal, and a new sample is required for each pulse (a so-called “diffract-and-destroy” approach) [78]. Multi-crystal sample delivery methods are therefore integral to these studies. The numerous techniques that have become established in recent years can be separated into two main categories, viz. injection and fixed-target methods.

Injection methods inject microcrystals suspended in a carrier medium, which in many cases is the crystallisation liquor, into the path of X-ray beam. A variety of injection systems have been developed [79, 80]. While relatively simple, this technique suffers from a low hit rate – i.e. most X-ray pulses will fail to hit a

crystal – and requires large amounts of material, which limits its practical application. This method also leads to unavoidable shot-to-shot variation in both the diffraction quality and power, which must be accounted for by scaling and merging during processing. For photocrystallographic studies in particular, given the strong dependence on the size and morphology of the crystal, one would also expect considerable variation in photoconversion level.

Fixed-target methods instead attach crystals to a target on a translation stage that allows the X-ray beam to scan the surface during the experiment. Mehrabi et al. developed a system specifically for TR studies in which 30 mm^2 patterned silicon chips, each containing 20,736 sample wells, are mounted on a nano-translation stage that enables fast raster scanning across the chip surface to expose different crystals to the X-ray beam [81]. The chip is usually also scanned prior to the experiment to identify which grid cells contain crystals, thereby removing an element of uncertainty and drastically improving the hit rate compared to injection methods. However, it is important to ensure during mounting that crystals are well separated to avoid contamination between measurements. The better control of the crystal sampling afforded by fixed-target methods makes them potentially better suited to chemical crystallography in general and photocrystallography in particular.

4.6 Data Processing

Another important consideration in a TR-SCXRD experiment is the data processing. Whereas most conventional spectroscopic techniques generate data that is straightforward to visualise – if not to interpret – the raw data collected in a SCXRD experiment requires extensive processing to analyse.

At a basic level, a traditional SCXRD dataset consists of a series of 2D diffraction patterns captured as the crystal is rotated to bring different regions of its reciprocal space onto the detector. The series of images is first subject to a peak-finding procedure, by which all of the reflections are identified, and the reflections are then indexed to obtain a unit cell, space group and an orientation matrix describing the orientation of the crystal with respect to the diffractometer axes. Once this is complete, the diffraction images are integrated to obtain the intensities of the diffraction spots, resulting in a list of relative intensities with assigned Miller indices (*hkl*). This information is then used to obtain an initial crystallographic model – a unit cell and a set of atom positions – in a structure solution, and the model is iteratively refined to obtain a best fit to the data.

There are a number of commercial and open-source software programs for SCXRD data processing. The vast majority have been developed for laboratory experiments and as such use graphical user interfaces (GUIs) that require considerable user interaction in their standard operation, although some programs offer automatic processing routes that aim to automate part or all of the workflow. As TR-SCXRD experiments typically produce vast amounts of data very rapidly by design, it is desirable, even essential, to obtain an initial set of solutions and

refinements "on-the-fly" to quickly confirm the quality of the measurements and to ascertain whether a chosen set of experimental conditions is giving the desired result. Such on-the-fly processing requires a fast, reliable auto-processing pipeline. We venture that this can be best achieved using command-line-based programs that can be controlled via an external script without the need for user interaction. Automated processing is particularly essential for making efficient use of the limited experiment time available at synchrotron and XFEL facilities; processing scripts can easily be incorporated into workflows so that they are started automatically and can provide real-time feedback to users to enable fast decision-making.

In a traditional pump-probe experiment, data can be simply accumulated until enough signal is obtained over an angular rotation, and once a full set of crystal orientations have been measured, the data can be processed immediately using standard workflows. For pump-multiprobe experiments, which record data for several time delays simultaneously, additional preprocessing steps are required to prepare the data for analysis. The data from individual time delays must be separated, and multiple images collected for a given time delay/rotation may then need to be summed to obtain sufficient signal-to-noise ratio. Automated routines are required to prevent this from becoming a bottleneck during an experiment, and given the huge amount of data involved, this may require access to high-performance computing clusters. After the data has been sorted and preprocessed, individual datasets are obtained for each time delay, which can be treated as regular SCXRD datasets and processed as outlined above.

Although ab initio structure solution may be possible with data from individual time delays, for rapid auto-processing routines, a good-quality starting model for both the ground state and the photostationary excited state is optimal. These can be obtained from low-temperature steady-state photocrystallographic measurements carried out in preparation for the TR-SCXRD study. By using these as reference models when refining the TR data, the occupations of known excited-state species can be quickly and automatically refined. Identifying unknown short-lived species is more complicated, but can be approached by looking for new features in Fourier electron density difference maps generated between the TR structures and the ground-state structural model (termed "photodifference maps"). If new short-lived species are identified, they are likely to be present at a low population level and will require a new disorder model to be created to refine the data against. At present there is no substitute for an experienced crystallographer in this situation, and user intervention is thus still required.

In summary, the aim of automatic data processing pipelines in pump-(multi)probe SCXRD experiments should be to automate the routine parts of the data processing and enable users to quickly identify datasets that show changes or new structural features of interest so that these can be investigated manually by an experienced crystallographer. Automatic processing pipelines are absolutely critical to the success of TR studies and will only become more important as these methods progress towards faster data collections at higher time resolution.

4.7 Sub-second Linkage Isomer Studies

Solid-state linkage isomer systems have traditionally been studied for their ability to access long-lived metastable states at low temperatures, and as such comparatively little work has been done to explore the potential for fast single-crystal switching at higher temperatures. However, the solid-state kinetics discussed in Sect. 3 highlight the strong temperature dependence of the excited-state lifetime, and spectroscopic studies conducted at or near ambient temperature indicate that faster pump-probe diffraction studies should be possible if the correct experiment temperature is selected [25].

In 2017, Casaretto et al. reported the first laboratory-based pump-probe diffraction study following the photoisomerisation in crystals of the archetypal iron-nitrosyl system sodium nitroprusside [82]. The laboratory pump-probe setup incorporated an electronically gated HPC detector, a microfocus X-ray source and a complementary transient absorption spectroscopy setup. The study followed changes in the relative intensity of Bragg reflections known to change significantly through the isomerisation, which were used to infer a conversion of *c.a.* 1% to a short-lived photo-induced isomer with 6 ms time resolution at 150 K. The results of this study are encouraging and pave the way towards further pump-probe diffraction studies on linkage isomer crystals, aiming to obtain full 3D crystal structures in line with the fast time-resolved experiments on other photoactive materials conducted at synchrotron and XFEL facilities.

5 Conclusions

This chapter has outlined some of the key developments and challenges in the study of solid-state linkage isomerism as a model for solid-state photochemical reactions and in particular has highlighted the wealth of structural and kinetic information that can be obtained by in situ single-crystal photocrystallography experiments. While long-lived metastable isomers are easily characterised using traditional X-ray photocrystallography methods in the laboratory, identifying shorter-lived excited-state species requires time-resolved experiments at synchrotron sources. With the increasing availability of state-of-the-art X-ray sources and detectors, it is conceivable that time-resolved photocrystallography measurements at millisecond and possibly even microsecond timescales may soon be routinely achievable in the laboratory. This will greatly support the development of faster time-resolved diffraction methods at synchrotron and XFEL facilities, helping to transform what is currently a specialist research area into a routine and widely available technique for real-time visualisation of solid-state photochemical reactions. This is an exciting prospect with the potential to revolutionise our understanding of photoactivated processes in the solid state and to drive the development of new and improved functional materials for applications including catalysis and molecular data storage.

References

1. Carducci MD, Pressprich MR, Coppens P (1997) Diffraction studies of photoexcited crystals: metastable nitrosyl-linkage isomers of sodium nitroprusside. J Am Chem Soc 119:2669–2678
2. Šrajer V, Teng T-Y, Ursby T, Pradervand C, Ren Z, Adachi S-I et al (1996) Photolysis of the carbon monoxide complex of myoglobin: nanosecond time- resolved crystallography. Science 274:1726–1729
3. Penland RB, Lane TJ, Quagliano JV (1956) Infrared absorption spectra of inorganic coördination complexes. VII. Structural isomerism of nitro- and nitritopentamminecobalt(III) chlorides1a,b. J Am Chem Soc 78:887–889
4. Murmann RK, Taube H (1956) The mechanism of the formation and rearrangement of nitritocobalt(III) ammines1. J Am Chem Soc 78:4886–4890
5. Fraser RTM (1967) Linkage isomerism. Werner centennial. In: Advances in chemistry, vol 62. American Chemical Society, pp 295–305
6. Chattopadhyay T, Ghosh M, Majee A, Nethaji M, Das D (2005) Linkage isomerism in 4-(2-aminoethyl)morpholine (L) complexes of nickel (II) nitrite: X-ray single crystal structure of trans-[$NiL_2(NO_2)_2$]. Polyhedron 24:1677–1681
7. Adell B (1944) Über die Geschwindigkeit der Umwandlung von Nitrito- in Nitropentamminkobalt(III)-chlorid. Zeitschrift für Anorganische Chem 252:272–280
8. Cohen MD, Schmidt GMJ (1964) 383. Topochemistry. Part I. a survey. J Chem Soc:1996–2000
9. Cohen MD, Schmidt GMJ, Sonntag FI (1964) 384. Topochemistry. Part II. The photochemistry of trans-cinnamic acids. J Chem Soc (Resumed):2000–2013
10. Ohashi Y (2013) Dynamic motion and various reaction paths of cobaloxime complexes in crystalline-state photoreaction. Crystallogr Rev 19:2–146
11. Ohashi Y, Sasada Y (1977) X-ray analysis of co-C bond cleavage in the crystalline state. Nature 267:142–144
12. Warren M, Brayshaw S, Johnson A, Schiffers S, Raithby P, Easun T et al (2009) Reversible 100% linkage isomerization in a single crystal to single crystal transformation: photocrystallographic identification of the metastable [Ni(dppe)(η^1-ONO)Cl] isomer. Angew Chem 121:5821–5824
13. Warren MR, Easun TL, Brayshaw SK, Deeth RJ, George MW, Johnson AL et al (2014) Solid-state interconversions: unique 100% reversible transformations between the ground and metastable states in single-crystals of a series of nickel(II) nitro complexes. Chem Eur J 20:5468–5477
14. Hatcher LE, Warren MR, Allan DR, Brayshaw SK, Johnson AL, Fuertes S et al (2011) Metastable linkage isomerism in [Ni(Et_4dien)(NO_2)$_2$]: a combined thermal and photocrystallographic structural investigation of a nitro/nitrito interconversion. Angew Chem Int Ed 50:8371–8374
15. Hatcher LE (2016) Raising the (metastable) bar: 100% photo-switching in [Pd(Bu_4dien)(η^1-$\underline{NO_2}$)]$^+$ approaches ambient temperature. CrystEngComm 18:4180–4187
16. Hatcher LE, Raithby PR (2017) The impact of hydrogen bonding on 100% photo-switching in solid-state nitro-nitrito linkage isomers. CrystEngComm 19:6297–6304
17. Hatcher LE, Bigos EJ, Bryant MJ, MacCready EM, Robinson TP, Saunders LK et al (2014) Thermal and photochemical control of nitro-nitrito linkage isomerism in single-crystals of [Ni(medpt)(NO_2)(η^2-ONO)]. CrystEngComm 16:8263–8271
18. Sylvester SO, Cole JM (2013) Quantifying crystallographically independent optical switching dynamics in Ru-SO_2 photoisomers via lock-and-key crystalline environment. J Phys Chem Lett 4:3221–3226
19. Cormary B, Ladeira S, Jacob K, Lacroix PG, Woike T, Schaniel D et al (2012) Structural influence on the photochromic response of a series of ruthenium mononitrosyl complexes. Inorg Chem 51:7492–7501

20. Warren MR, Brayshaw SK, Hatcher LE, Johnson AL, Schiffers S, Warren AJ et al (2012) Photoactivated linkage isomerism in single crystals of nickel, palladium and platinum di-nitro complexes – a photocrystallographic investigation. Dalton Trans 41:13173–13179
21. Atkins P, Atkins PW, Shriver DF (2006) Shriver & Atkins inorganic chemistry. W.H. Freeman
22. Schaniel D, Woike T (2009) Necessary conditions for the photogeneration of nitrosyl linkage isomers. Phys Chem Chem Phys 11:4391–4395
23. Hauser U, Oestreich V, Rohrweck HD (1977) On optical dispersion in transparent molecular systems. Zeitschrift für Phys A Hadrons Nucl 280:17–25
24. Woike T, Krasser W, Zöllner H, Kirchner W, Haussühl S (1993) Population dynamics of the two light induced metastable states in $Na_2[Fe(CN)_5NO]\cdot 2H_2O$ single crystals. Zeitschrift für Phys D Atoms Mol Clust 25:351–356
25. Schaniel D, Nicoul M, Woike T (2010) Ultrafast reversible ligand isomerisation in $Na_2[Fe(CN)_5NO]\cdot 2H_2O$ single crystals. Phys Chem Chem Phys 12:9029–9033
26. Coppens P, Novozhilova I, Kovalevsky A (2002) Photoinduced linkage isomers of transition-metal nitrosyl compounds and related complexes. Chem Rev 102:861–884
27. Schaniel D, Bendeif EE, Woike T, Böttcher HC, Pillet S (2018) Wavelength-selective photoisomerisation of nitric oxide and nitrite in a rhodium complex. CrystEngComm 20:7100–7108
28. Fomitchev DV, Furlani TR, Coppens P (1998) Combined X-ray diffraction and density functional study of $[Ni(NO)(\eta^5\text{-}Cp^*)]$ in the ground and light-induced metastable states. Inorg Chem 37:1519–1526
29. Fomitchev DV, Coppens P (1996) X-ray diffraction analysis of geometry changes upon excitation: the ground-state and metastable-state structures of $K_2[Ru(NO_2)_4(OH)(NO)]$. Inorg Chem 35:7021–7026
30. Fomitchev DV, Coppens P, Li T, Bagley KA, Chen L, Richter-Addo GB (1999) Photo-induced metastable linkage isomers of ruthenium nitrosyl porphyrins. Chem Commun:2013–2014
31. Schaniel D, Cormary B, Malfant I, Valade L, Woike T, Delley B et al (2007) Photogeneration of two metastable NO linkage isomers with high populations of up to 76% in trans-$[RuCl(py)_4(NO)][PF_6]_2\cdot \frac{1}{2}H_2O$. Phys Chem Chem Phys 9:3717–3724
32. Cormary B, Malfant I, Valade L, Buron-Le Cointe M, Toupet L, Todorova T et al (2009) $[Ru(py)_4Cl(NO)](PF_6)_2\cdot 0.5H_0O$: a model system for structural determination and ab initio calculations of photo-induced linkage NO isomers. Erratum. Acta Crystallogr B 65: 787–787
33. Kovalevsky AY, King G, Bagley KA, Coppens P (2005) Photoinduced oxygen transfer and double-linkage isomerism in a *cis*-$(NO)(NO_2)$ transition-metal complex by photocrystallography, FT-IR spectroscopy and DFT calculations. Chem Eur J 11:7254–7264
34. Amabilino S, Tasse M, Lacroix PG, Mallet-Ladeira S, Pimienta V, Akl J et al (2017) Photorelease of nitric oxide (NO) on ruthenium nitrosyl complexes with phenyl substituted terpyridines. New J Chem 41:7371–7383
35. Giglmeier H, Kerscher T, Klufers P, Schaniel D, Woike T (2009) Nitric-oxide photorelease and photoinduced linkage isomerism on solid $[Ru(NO)(terpy)(L)]BPh_4$ (L = glycolate dianion). Dalton Trans 2009:9113–9116
36. Dieckmann V, Imlau M, Taffa DH, Walder L, Lepski R, Schaniel D et al (2010) Phototriggered NO and CN release from $[Fe(CN)_5NO]^{2-}$ molecules electrostatically attached to TiO_2 surfaces. Phys Chem Chem Phys 12:3283–3288
37. Johnson DA, Dew VC (1979) Photochemical linkage isomerization in coordinated sulfur dioxide. Inorg Chem 18:3273–3274
38. Kovalevsky AY, Bagley KA, Cole JM, Coppens P (2002) Light-induced metastable linkage isomers of ruthenium sulfur dioxide complexes. Inorg Chem 42:140–147
39. Kovalevsky AY, Bagley KA, Coppens P (2002) The first photocrystallographic evidence for light-induced metastable linkage isomers of ruthenium sulfur dioxide complexes. J Am Chem Soc 124:9241–9248

40. Bowes KF, Cole JM, Husheer SLG, Raithby PR, Savarese TL, Sparkes HA et al (2006) Photocrystallographic structure determination of a new geometric isomer of $[Ru(NH_3)_4(H_2O)(\eta^1\text{-OSO})]$ $[MeC_6H_4SO_3]_2$. Chem Commun:2448–2450
41. Phillips AE, Cole JM, d'Almeida T, Low KS (2010) Effects of the reaction cavity on metastable optical excitation in ruthenium-sulfur dioxide complexes. Phys Rev B 82:155118
42. Phillips AE, Cole JM, d'Almeida T, Low KS (2012) Ru–OSO coordination photogenerated at 100 K in tetraammineaqua(sulfur dioxide)ruthenium(II) (±)-camphorsulfonate. Inorg Chem 51:1204–1206
43. Sylvester SO, Cole JM, Waddell PG (2012) Photoconversion bonding mechanism in ruthenium sulfur dioxide linkage photoisomers revealed by in situ diffraction. J Am Chem Soc 134:11860–11863
44. Cole JM, Velazquez-Garcia JJ, Gosztola DJ, Wang SG, Chen Y-S (2018) η^2-SO_2 linkage photoisomer of an osmium coordination complex. Inorg Chem 57:2673–2677
45. Allen AD, Senoff CV (1965) Nitrogenopentammineruthenium(II) complexes. Chem Commun (Lond):621–622
46. Tanaka H, Nishibayashi Y, Yoshizawa K (2016) Interplay between theory and experiment for ammonia synthesis catalyzed by transition metal complexes. Acc Chem Res 49:987–995
47. Nishibayashi Y (2018) Development of catalytic nitrogen fixation using transition metal-dinitrogen complexes under mild reaction conditions. Dalton Trans 47:11290–11297
48. Yang Y, Liu J, Wei Z, Wang S, Ma J (2019) Transition metal-dinitrogen complex embedded graphene for nitrogen reduction reaction. ChemCatChem 11:2821–2827
49. Fomitchev DV, Bagley KA, Coppens P (2000) The first crystallographic evidence for side-on coordination of N_2 to a single metal center in a photoinduced metastable state. J Am Chem Soc 122:532–533
50. Goulkov M, Schaniel D, Woike T (2010) Pulse recording of thermal and linkage isomer gratings in nitrosyl compounds. J Opt Soc Am B 27:927–932
51. Sylvester SO, Cole JM (2013) Solar-powered nanomechanical transduction from crystalline molecular rotors. Adv Mater 25:3324–3328
52. Cole JM, Yeung KYM, Pace G, Sylvester SO, Mersch D, Friend RH (2015) In situ synthesis, crystallisation, and thin-film processing of single crystals of *trans*-$[Ru(SO_2)(NH_3)_4(H_2O)][p\text{-}TolSO_3]_2$ bearing SO_2 linkage photo-isomers: towards optical device applications. CrystEngComm 17:5026–5031
53. Brayshaw SK, Knight JW, Raithby PR, Savarese TL, Schiffers S, Teat SJ et al (2010) Photocrystallography – design and methodology for the use of a light-emitting diode device. J Appl Crystallogr 43:337–340
54. Kaminski R, Jarzembska KN, Kutyla SE, Kaminski M (2016) A portable light-delivery device for in situ photocrystallographic experiments in the home laboratory. J Appl Crystallogr 49:1383–1387
55. Avrami M (1941) Kinetics of phase change. III. Granulation, phase change, and microstructure. J Chem Phys 9:177–184
56. Avrami M (1940) Kinetics of phase change. II. Transformation-time relations for random distribution of nuclei. J Chem Phys 8:212–224
57. Avrami M (1939) Kinetics of phase change. I. General theory. J Chem Phys 7:1103–1112
58. Hatcher LE, Christensen J, Hamilton ML, Trincao J, Allan DR, Warren MR et al (2014) Steady-state and pseudo-steady-state photocrystallographic studies on linkage isomers of $[Ni(Et_4dien)(\eta^2\text{-O,ON})(\eta^1\text{-}NO_2)]$: identification of a new linkage isomer. Chem Eur J 20:3128–3134
59. Hatcher LE, Skelton JM, Warren MR, Stubbs C, da Silva EL, Raithby PR (2018) Monitoring photo-induced population dynamics in metastable linkage isomer crystals: a crystallographic kinetic study of $[Pd(Bu_4dien)NO_2]BPh_4$. Phys Chem Chem Phys 20:5874–5886
60. Benedict JB, Coppens P (2009) Kinetics of the single-crystal to single-crystal two-photon photodimerization of α-trans-cinnamic acid to α-truxillic acid. J Phys Chem A 113:3116–3120

61. Bertmer M, Nieuwendaal RC, Barnes AB, Hayes SE (2006) Solid-state photodimerization kinetics of α-trans-cinnamic acid to α-truxillic acid studied via solid-state NMR. J Phys Chem B 110:6270–6273
62. Abdelmoty I, Buchholz V, Di L, Guo C, Kowitz K, Enkelmann V et al (2005) Polymorphism of cinnamic and α-truxillic acids: new additions to an old story. Cryst Growth Des 5:2210–2217
63. Enkelmann V, Wegner G, Novak K, Wagener KB (1993) Single-crystal-to-single-crystal photodimerization of cinnamic acid. J Am Chem Soc 115:10390–10391
64. More R, Busse G, Hallmann J, Paulmann C, Scholz M, Techert S (2010) Photodimerization of crystalline 9-anthracenecarboxylic acid: a nontopotactic autocatalytic transformation. J Phys Chem C 114:4142–4148
65. Skelton JM, Crespo-Otero R, Hatcher LE, Parker SC, Raithby PR, Walsh A (2015) Energetics, thermal isomerisation and photochemistry of the linkage-isomer system [Ni(Et_4dien)(η^2-O,ON)(η^1-NO_2)]. CrystEngComm 17:383–394
66. Grünbein ML, Stricker M, Nass Kovacs G, Kloos M, Doak RB, Shoeman RL et al (2020) Illumination guidelines for ultrafast pump–probe experiments by serial femtosecond crystallography. Nat Methods 17:681–684
67. Skarzynski T (2013) Collecting data in the home laboratory: evolution of X-ray sources, detectors and working practices. Acta Crystallogr D Biol Crystallogr 69:1283–1288
68. Bilderback DH, Elleaume P, Weckert E (2005) Review of third and next generation synchrotron light sources. J Phys B Atomic Mol Phys 38:S773–S797
69. Jaeschke EJ, Khan S, Schneider JR, Hastings JB (2016) Synchrotron light sources and free-electron lasers: accelerator physics, instrumentation and science applications. Springer, Berlin
70. Geloni G, Huang Z, Pellegrini C (2017) Chapter 1 the physics and status of X-ray free-electron lasers. In: X-ray free electron lasers: applications in materials, chemistry and biology. The Royal Society of Chemistry, pp 1–44
71. Mobilio S, Boscherini F, Meneghini C (2016) Synchrotron radiation: basics, methods and applications. Springer, Berlin
72. Allé P, Wenger E, Dahaoui S, Schaniel D, Lecomte C (2016) Comparison of CCD, CMOS and hybrid pixel x-ray detectors: detection principle and data quality. Phys Scr 91:063001
73. He T, Durst R, Becker B, Kaercher J, Wachter G (2011) A large area X-ray imager with online linearization and noise suppression. SPIE
74. Förster A, Brandstetter S, Schulze-Briese C (2019) Transforming X-ray detection with hybrid photon counting detectors. Philos Trans R Soc A Math Phys Eng Sci 377:20180241
75. Blaj G, Caragiulo P, Carini G, Carron S, Dragone A, Freytag D et al (2015) X-ray detectors at the Linac coherent light source. J Synchrotron Radiat 22:577–583
76. Bergamaschi A, Mozzanica A, Schmitt B (2020) XFEL detectors. Nat Rev Phys 2:335–336
77. Poikela T, Plosila J, Westerlund T, Campbell M, Gaspari MD, Llopart X et al (2014) Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout. J Instrum 9:C05013–C05013
78. Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A et al (2011) Femtosecond X-ray protein nanocrystallography. Nature 470:73–77
79. Grunbein ML, Nass KG (2019) Sample delivery for serial crystallography at free-electron lasers and synchrotrons. Acta Crystallogr Sect D 75:178–191
80. Martiel I, Muller-Werkmeister HM, Cohen AE (2019) Strategies for sample delivery for femtosecond crystallography. Acta Crystallogr Sect D 75:160–177
81. Mehrabi P, Muller-Werkmeister HM, Leimkohl J-P, Schikora H, Ninkovic J, Krivokuca S et al (2020) The HARE chip for efficient time-resolved serial synchrotron crystallography. J Synchrotron Radiat 27:360–370
82. Casaretto N, Schaniel D, Alle P, Wenger E, Parois P, Fournier B et al (2017) In-house time-resolved photocrystallography on the millisecond timescale using a gated X-ray hybrid pixel area detector. Acta Crystallogr B 73:696–707

Struct Bond (2020) 185: 239–272
https://doi.org/10.1007/430_2020_74

Published online: 4 October 2020

Time-Resolved Single-Crystal X-Ray Crystallography

Paul R. Raithby

Contents

Abstract In this chapter the development of time-resolved crystallography is traced from its beginnings more than 30 years ago. The importance of being able to "watch" chemical processes as they occur rather than just being limited to three-dimensional pictures of the reactant and final product is emphasised, and time-resolved crystallography provides the opportunity to bring the dimension of time into the crystallographic experiment. The technique has evolved in time with developments in technology: synchrotron radiation, cryoscopic techniques, tuneable lasers, increased computing power and vastly improved X-ray detectors. The shorter the lifetime of the species being studied, the more complex is the experiment. The chapter focusses on the results of solid-state reactions that are activated by light, since this process does not require the addition of a reagent to the crystalline material and the single-crystalline nature of the solid may be preserved. Because of this photoactivation, time-resolved crystallography is often described as "photocrystallography".

P. R. Raithby (✉)
Department of Chemistry, University of Bath, Bath, UK
e-mail: p.r.raithby@bath.ac.uk

The initial photocrystallographic studies were carried out on molecular complexes that either underwent irreversible photoactivated processes where the conversion took hours or days. Structural snapshots were taken during the process. Materials that achieved a metastable state under photoactivation and the excited (metastable) state had a long enough lifetime for the data from the crystal to be collected and the structure solved. For systems with shorter lifetimes, the first time-resolved results were obtained for macromolecular structures, where pulsed lasers were used to pump up the short lifetime excited state species and their structures were probed by using synchronised X-ray pulses from a high-intensity source. Developments in molecular crystallography soon followed, initially with monochromatic X-ray radiation, and pump-probe techniques were used to establish the structures of photoactivated molecules with lifetimes in the micro- to millisecond range. For molecules with even shorter lifetimes in the sub-microsecond range, Laue diffraction methods (rather than using monochromatic radiation) were employed to speed up the data collections and reduce crystal damage. Future developments in time-resolved crystallography are likely to involve the use of XFELs to complete "single-shot" time-resolved diffraction studies that are already proving successful in the macromolecular crystallographic field.

Keywords Excited state lifetimes · Lasers · Macromolecules · Metastable states · Photochemistry · Photocrystallography · Solid-state · Synchrotron radiation · Time-resolved crystallography · XFELs · X-rays

Abbreviations

CCD	Charge coupled device
DFT	Density functional theory
ES	Excited state
ESRF	European Synchrotron Radiation Facility
EXAFS	Extended X-ray absorption fine structure
FWHM	Full width at half maximum
GS	Ground state
HATRX	Hadamard time-resolved crystallographic experiment
LIESST	Light-induced excited spin-state trapping experiment
PES	Photoelectron spectrum
RATIO	Method for collecting sequential laser-*on* and laser-*off* X-ray data
XAFS	X-ray absorption fine structure
XFEL	X-ray free-electron laser

1 Introduction

Over the last century, single-crystal X-ray crystallography has developed into the optimum method for determining the molecular structure of materials in the crystalline state and, as this volume of Structure and Bonding shows, now underpins many aspects of the physical and life sciences. What the standard crystallographic method provides is a full three-dimensional picture of the structure of the starting material and of the product of the reaction if both can be obtained in a crystalline form. What it does not do is provide a pathway by which the starting material is converted into the product, and, since the reaction may not occur in the solid state, the immediate relevance to a solution or gas phase chemical reaction may not be apparent [1].

The reason for this inability to follow a solid-state chemical process is because the single-crystal X-ray crystallographic experiment is both space and time averaged. In terms of space averaging, every unit cell in the crystal contributes to the diffraction pattern obtained, so if some molecules within the crystal are changing and others are not, an average of the structures will be obtained. This is commonly observed in crystal structures when, for example, various parts of the molecule adopt different orientations in different unit cells or lattice solvent molecules adopt different orientations; this is termed disorder [2]. The crystallographic experiments are also time averaged. Although a single X-ray photon may interact with the electron cloud that surrounds an atomic nucleus in crystal in 10^{-18} s, it has to be remembered that in a single-crystal X-ray crystallographic experiment, the whole crystal has to be sampled. Typically, the crystal may contain 10^{15} molecules or more, and even with modern laboratory-based diffractometers, the sampling process takes between minutes and hours, so that the crystal structure obtained is an average over the whole duration of the data collection [3].

A further problem associated with following chemical or biological processes within a single crystal is the retention of crystallinity throughout the process since a loss of crystallinity caused by degradation of the crystal means that a single-crystal diffraction experiment is no longer viable. This limits the type of reaction that can be followed. Adding a reagent to a crystal to facilitate a chemical reaction will usually destroy the crystal, although there are an increasing number of examples, often associated with crystals of framework structure materials, where small molecules (liquids or gases) can be introduced into the crystal and can undergo a reversible or irreversible physical or chemical process [4–6]. A much easier way of facilitating a single-crystal-to-single-crystal reaction is to use an external medium such as irradiation with light [7–9], application of pressure [10–12] or change in temperature [13–15] which is less likely to disrupt the crystalline lattice, and such experiments have been becoming more feasible over the last few decades because of advances in technology.

Over the last three decades, developments in synchrotron facilities, X-ray detectors, cryogenic apparatus, laser technology, computing power and data storage capacity have all enabled single-crystal X-ray studies to provide information about solid-state reactions and dynamic processes that occur in crystals. Now chemical and

biological processes that occur in crystals can be monitored in "real time", and three-dimensional structural information on species that exist in the solid state for only fractions of a second may be obtained [16–18]. Regarding the external media for inducing reactions in crystals, the use of "light" has proved the most popular and is arguably the least likely to cause crystal degradation. Photochemically induced chemical reactions are relatively easy to control. The size, shape and intensity of the incident beam can be controlled using a wide variety of optical equipment and, when using a tuneable source, the wavelength can be varied to match the process involved. This has led to the development of "photocrystallography" [19], a term attributed to Philip Coppens [20], one of the pioneers in the area, although the crystallography of photoactivated species is also included in the description of "photosalient" processes [21] or, more generally, as "time-resolved" crystallography. With all the technological developments, the speed of a full single-crystal X-ray data collection has been reduced to minutes particularly when a high-intensity synchrotron source is used. This has provided the opportunity to undertake a whole range of new crystallographic experiments that can follow the dynamics of a chemical process and obtain a full three-dimensional picture of photoactivated species. This chapter will highlight some of the technological and experimental results that have been obtained in the first 20 years of the twenty-first century.

Of all the technological advances that have facilitated the growth of photocrystallography, it is the more general availability of synchrotrons as research tools that has proved to be most important. They have been used to apply a range of analytical techniques to study chemical and biological samples by providing high-intensity electromagnetic radiation as the probe. A synchrotron is a storage ring in which electrons are accelerated around the ring at speeds approaching the speed of light and, as a consequence, are subject to relativistic effects. This causes the electrons to circulate around the ring in discrete bunches. The resultant radiation is used in different experimental stations which are at the end of beamlines radiating out tangentially from the storage ring. Some of these beamlines are dedicated to crystallographic experiments and use X-rays with wavelengths in the range 0.4–2.5 Å [22]. A schematic representation of a synchrotron ring is illustrated in Fig. 1 which shows how the electrons are initially accelerated from their source through a linear accelerator, through a booster synchrotron into the main storage ring where their circular orbits are controlled by magnetic fields and their energies are controlled by insertion devices which guide the X-ray beam to the experimental hutch where the X-ray diffractometer sits.

2 Photocrystallographic Methodology

The type of photocrystallographic experiment that can be carried out is dependent on the lifetime of the photoactivated species that is being studied. The lifetimes of the various processes can range from hours to femtoseconds as are summarised in Fig. 2 so that the type of experiment that can be carried out and the results that can be

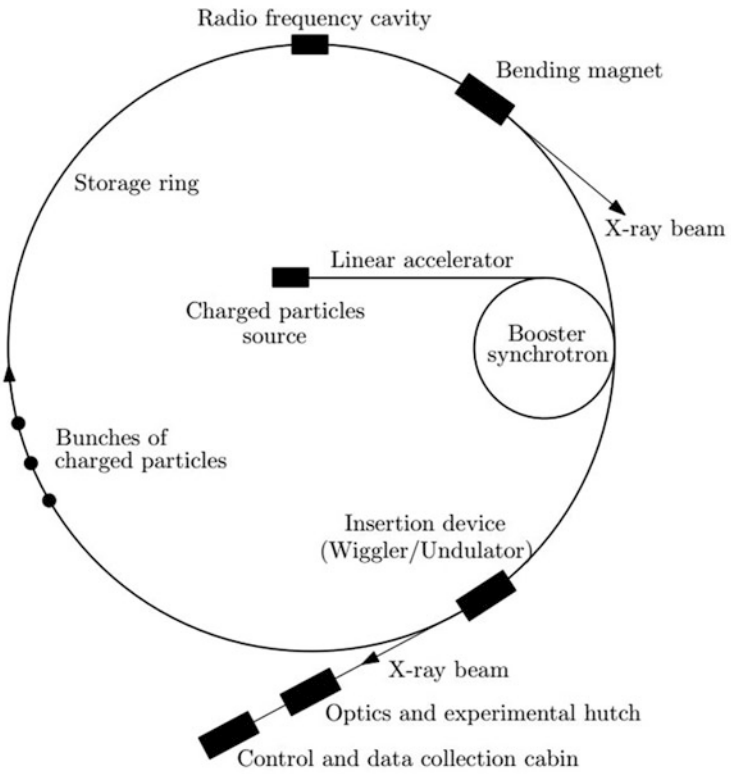


Fig. 1 A schematic of a synchrotron ring showing a tangential beamline

obtained must be considered carefully beforehand. The photoactivated "excited state" (ES) must exist for longer than the time required for the data collection, although it can be maintained by repeatedly or continuously pumping light into the system. The experiment and the chemical process that is being studied must be brought on to a common timescale for the experiment to be successful. This can be achieved either by slowing the photoreaction down to the timeframe of the crystallographic experiment or speeding up the experiment to match the lifetime of the excited state species. The photoreaction can be slowed down by effectively using trapping strategies, e.g. chemical- or cryo-trapping methods, which involve a sudden change in the reaction conditions in order to "freeze" the reactant in a transient state for a period of time long enough to permit the analysis. While there are some advantages to this approach, the trapping process may change the natural progress of the solid-state reaction. To be sure of observing the true reaction pathway, it may be advisable to adopt the second strategy and speed up the data collection methodology so that the photoinduced process can be followed in real time.

In summary, the shorter the lifetime of the photoexcited species, the more challenging is the photocrystallographic experiment that is required to characterise

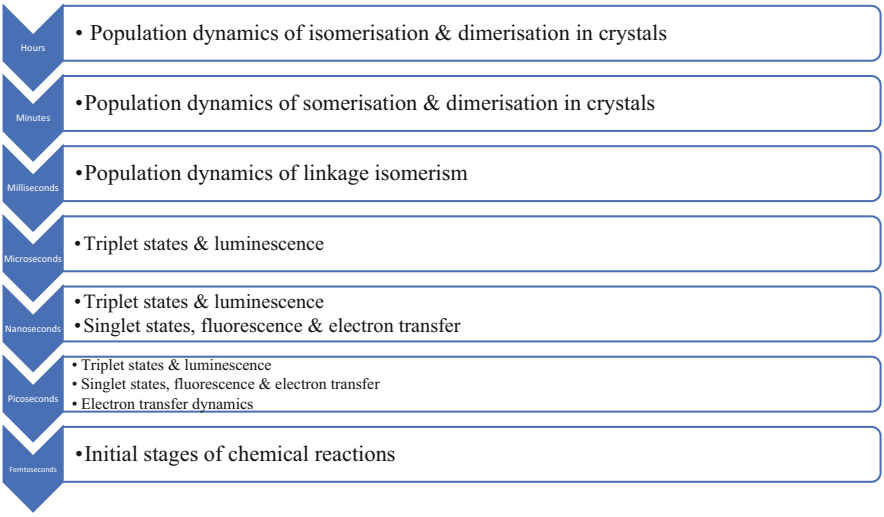


Fig. 2 Timescales of the dynamic processes that occur in chemistry

the structure of the excited state. Photocrystallographic experiments can be characterised into three types depending on the lifetime of the excited state species.

2.1 Steady-State and Pseudo-Steady-State Methodologies

At longer timescales, from milliseconds upwards, "steady-state" or "pseudo-steady-state" experiments can be carried out using standard single-crystal X-ray diffraction methods [8, 18, 23] with monochromated X-ray radiation. The steady-state methodology is used typically to study metastable excited states, those with lifetimes of hours up to infinity, if the excited state is generated and maintained at the appropriate low temperature [24]. The metastable state is generated, at a given temperature, by irradiating the sample for a period long enough to maximise the conversion from the ground state. The irradiation is then stopped, and a standard single-crystal X-ray data collection is performed. Under these experimental conditions, there are no concerns about sample heating from the irradiation source since this has been switched off prior to the start of the data collection. This process is illustrated in Fig. 3a where the ground state, unexcited structure is collected first (in order to provide a benchmark against which changes in the photogenerated excited state can be compared), then the excited state is generated with light irradiation and finally the structure is redetermined using X-rays, the light source having been turned off. Pseudo-steady-state methodology is used to study samples with slightly shorter excited state lifetimes, usually in the range of milliseconds to minutes. In these experiments the

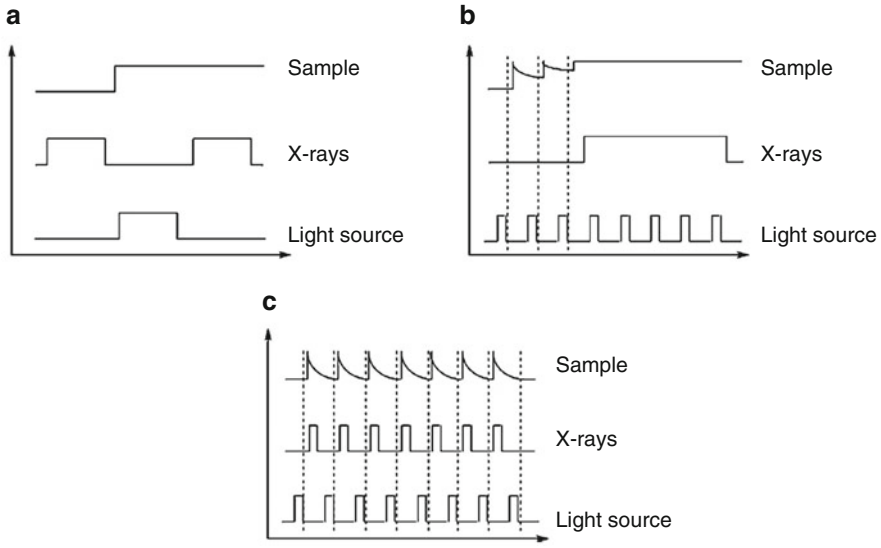


Fig. 3 Schematic diagrams to show the timing sequences used in the different types of photocrystallographic experiments: (**a**) steady-state methods, (**b**) pseudo-steady-state methods and (**c**) pump-probe methods. *Taken from Ref.* [18] *with permission from Elsevier*

crystalline sample is irradiated throughout the data collection to maintain a constant excited state occupancy, as illustrated in Fig. 3b. This shows the excited state being pumped up to a "steady state" with repeated pulses of light before the X-ray data collection commences, the excited state being maintained throughout the data collection by continued pulses of light. Thus, this method requires an effective means of illuminating the crystal fully throughout the data collection without the movement of the diffractometer blocking the illumination source at any point. Since the sample is irradiated throughout the experiment, heating effects at the sample resulting from the irradiation source may result and have to be mitigated.

2.2 Stroboscopic or Pump-Probe Methodologies

In order to study transient photoinduced species, with lifetimes of microseconds and below, "stroboscopic" or "pump-probe" photocrystallographic methods are required as shown in Fig. 3c, the sample being repeatedly re-excited and only probed by the X-rays when in the excited state. The experiments generally require short-duration light and monochromatic X-ray pulses to be generated that are synchronised to arrive at the sample position in a specific time sequence [25] or to use a time-gated detector which is synchronised with the light pulses [26] so that X-ray data is only recorded when the crystal is activated. The light source for these experiments is usually a

pulsed laser that generates pulses on the nanosecond or picosecond timescale. The X-ray source can be pulsed in a number of ways. Initially, for experiments involving crystalline species with lifetimes in the nanosecond to microsecond range, the X-ray pulses were generated by placing a mechanical chopper in the incident beam [25, 27] which interrupts the beam so that the X-rays are only "on" in synchronisation with the laser pulse. When using these short pulses, the X-ray flux that impinges on the crystal is limited, and in order to overcome this problem, high-intensity synchrotron radiation is normally required. Even with this higher X-ray intensity, many pump-probe laser and X-ray cycles are required per data collection frame to build up a sufficiently strong diffraction image, and many frames are required in order to obtain a complete X-ray data set. The repeated pumping and probing usually has a detrimental effect on the crystalline sample, and significant sample heating effects may also be a problem.

One development to overcome the crystal degradation problem has been devised and used effectively in the macromolecular crystallographic community, and that is the pump-multi-probe method based on the Hadamard transformation [28]. In the conventional pump-probe approach described above, the photoactivation event is initiated by a laser pulse and then probed at a later time by an X-ray pulse, so that every laser (pump) pulse is followed by a single probe (X-ray) pulse after a predetermined time delay (see Fig. 4a). Therefore, to measure n time points, n pump-probe pairs are required. In contrast, in the Hadamard approach, each pump pulse is followed by a sequence of probe pulses, and the total signal from each sequence is recorded in a single measurement (see Fig. 4b). The sensitivity of the experiment is thus defined by the total number of photons in the complete probe sequence, with the time resolution defined as the total probe sequence length divided by the number of pulses. This method no longer limits the time resolution that can be achieved to the brilliance of the X-ray source by summing the time points across the probe sequence. It also gives an improved signal-to-noise ratio because of the larger number of photons recorded during the measurement.

As in standard pump-probe experiments, n pump-probe sequences are required to measure n time points. The pattern of the probe sequence is represented as rows of a nxn matrix (S) obtained from the Hadamard sequence. The simplest case is shown in Fig. 4b where each row of the matrix (and the probe sequence) is obtained by cycling by one element from the previous row to the left (see Fig. 4c).

For the Hadamard time-resolved experiment, the photoactivation is initiated and then the complete probe sequence (first row of the **S** matrix) is recorded as a single image. This process is repeated on the sample after relaxation but with the probe sequence now defined by the second row of the **S** matrix, until all the rows have been completed. The resulting encoded signals from the n excitations are then collated to form a vector **W** of length n. The time-dependent signal, I_t, is then obtained by reversing the probe sequence encoding by multiplying the vector W by the inverse of the matrix **S**, so that $I_t = \mathbf{S^{-1}W}$. A time-resolved crystallographic Hadamard experiment has been dubbed a HATRX experiment.

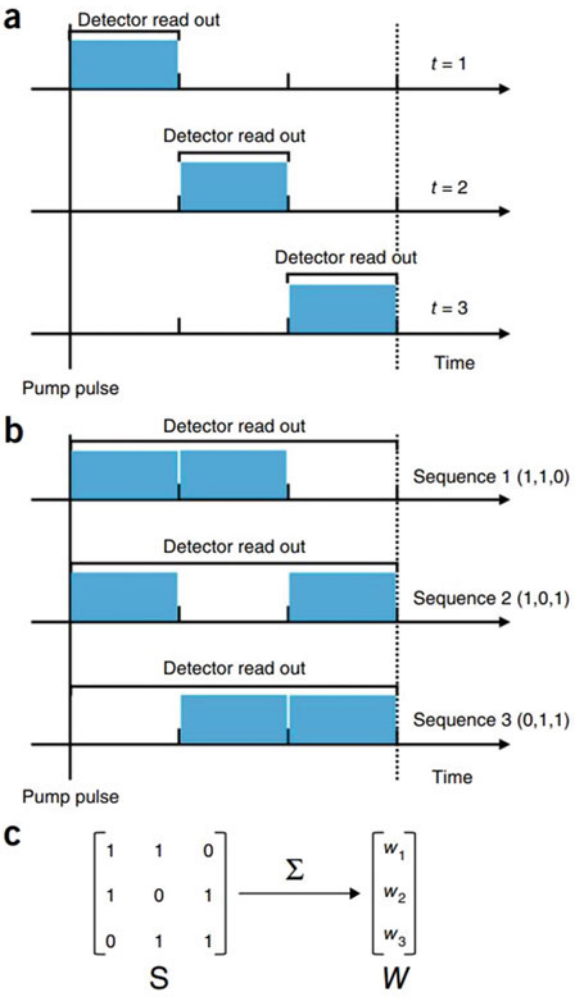


Fig. 4 (**a**) The time sequence for the classical pump-probe method showing three time delays. (**b**) The simplest Hadamard pulse sequence set up to measure three time points. (**c**) The 3 × 3 Hadamard **S** matrix showing how each row produces a single summed intensity for each reflection on the detector (w_1, . . ., w_n) forming the vector **W**. *Figure reproduced from Ref.* [28] *with permission from Nature Methods*

2.3 Laue Methods

The time-resolved experiments described so far have used monochromatic X-ray radiation. An alternative way of overcoming the limited flux in short-duration time-resolved crystallographic experiments is to change from monochromatic radiation to the Laue method where polychromatic X-ray radiation is used [29, 30]. The main advantage of the Laue method is the broader energy range used, compared with the narrow beam from a monochromator. This results in a substantial increase in intensity. The method used in both molecular [31] and macromolecular [32] crystallographic experiments is often called the pink-Laue method as "pink" indicates quite a small range of wavelengths as opposed to all the wavelengths available in the

full “white” beam. The pink-Laue method eliminates the need for the thousands of pump-probe cycles required in the stroboscopic method used for monochromatic data sets [33]. Because of the reduced time for the experiment, the crystal deterioration caused by laser and X-ray exposure may be reduced and the heating caused by the repeated laser pulses may also be reduced. In the best possible case, a single synchrotron pulse is sufficient to record a good enough diffraction pattern to solve the crystal structure [34]. In other examples several pump-probe cycles are required before a sufficiently intense diffraction pattern is obtained.

For pink-Laue data, a wavelength dependence correction has to be applied to the intensity data, because diffraction spots are being obtained from many wavelengths, and there also has to be a fitting of intensities of equivalent reflections that appear. With molecular crystals having relatively small unit cells, this scaling of intensities is challenging because there are relatively few of them. To avoid this difficulty, the RATIO method can be used [35], a method that will be explained in further detail below. A second complication with pink-Laue data is that there is a steep slope of wavelength distribution on the high-energy side of the pink-Laue spectrum. A small change in the unit cell dimensions on exposure, as would be expected with excitation, might lead to anomalous values of the *on/off* ratios of reflections scattered by wavelengths in this narrow region. The effect would be small if the conversion percentage from the ground to the excited state is small and also if the temperature increase is small. The affected reflections can be identified in the analysis of equivalent reflections and confirmed by checking the calculated wavelength from the Bragg angle and the *hkl* index after indexing. These can be removed from the averaging procedures and used to establish the structural changes that have taken place in the excited state.

As indicated above, the most sensitive measure of the structural change in a crystal when it is photoactivated into an excited state is the observed change in intensity of each of the reflections. In the RATIO method [35], this intensity difference is identified by using the *on/off* ratios as the observables in the activated structure refinement program LASER [36]. As explained in the previous paragraph, the advantage of the RATIO method when used in a Laue data collection is the elimination of the need to have a spectral curve to determine the wavelength at which each reflection is observed. To exploit the method, the laser-*on* and laser-*off* intensities for each reflection need to be collected immediately after one another. This eliminates variations in the intensity of the X-ray beam over time. This, in turn, adds an error to the intensity of the individual reflections and also eliminates the effect of any slow deterioration of the crystalline sample. In addition, slight differences in the absorption correction may occur if the laser-*on* and laser-*off* reflections were collected at different times in different settings. Finally, scaling is not required as the paired frames are collected at the same temperature and under the same conditions.

Even with these advances, it is often necessary to use multiple crystals of a crystalline sample, because of crystal deterioration, in order to obtain a complete data set. Scaling of the data sets is then required before electron density maps can be calculated and the structure solved and refined. Various scaling methods have been

applied from the simplest form which is based on the fractional change of each of the reflections on exposure to light to more sophisticated weighted least-squares scaling [31, 37].

Once the Laue data has been collected and the data set corrected, as with monochromatic photocrystallographic experiments, it is usual to compute a Fourier photodifference map to evaluate the structural changes that occur upon photoexcitation [38]. When using the RATIO method, the photodifference Fourier map is simply based on the difference between the observed laser-*on* and laser-*off* structure factors [29, 39]. These calculations are based on the fact that the crystal does not change phase (crystal system or space group) upon excitation which is generally true because percentage conversions from ground to excited state structures upon excitation are low for the short-lived species generated.

In an alternative approach to the pump-probe experiments, rather than synchronising the laser pump pulses with X-ray probe pulses generated by use of a mechanical chopper, new detector technology has allowed the X-ray detector to be synchronised with the laser pulse so that the detector only records X-ray intensity while the laser is on (or records after a designated time delay). These new detectors have fast read-out times, have the ability to internally stack series of recorded images and, most importantly, provide very fast and reliable "gating" affording fine control of when the detector is recording or not. With the efficiency of these detectors, some of the limitations of the X-ray flux can be minimised, and it is possible to consider the use of laboratory X-ray sources instead of synchrotrons for some longer lifetime photocrystallographic experiments. The use of a gated hybrid pixel detector, mounted on a conventional laboratory X-ray diffractometer, has been proven in an analysis of the photoinduced linkage isomerism of sodium nitroprusside [26]. The light-induced intensity variation between ground and excited states was detected at the 1% level, caused by the photoswitching of the nitrosyl group, and this change could be detected in a 6 microsecond window. The experimental approach is illustrated in Fig. 5. In the experiment a continuous X-ray beam impinges on the sample, and during the pump-probe cycle, with the structure continuously changing, the scattered diffraction pattern is sampled by the gated detector. The X-ray signal is only detected during a short adjustable window $X(t)$. The laser pulse serves as a trigger for the gating of the detector and synchronisation with a tuneable time delay Δt. The maximum time resolution possible is also dependent on the electronic response time of the X-ray detector.

With the gated hybrid pixel detector, the photon counting statistics determine the quality of the data, and there is no dark current or read-out noise as with a conventional CCD detector. The signal acquisition is defined by a tuneable measurement time window ($X(t)$ in Fig. 5) whose temporal width is only limited by the detector response time, which is the order of 100–200 ns. Also, with the pixel detector, it is possible to have a number of simultaneous measurement windows, with different delay times that can be acquired by the detector at the same time. This means that multiple time-resolved experiments could be carried out at the same time.

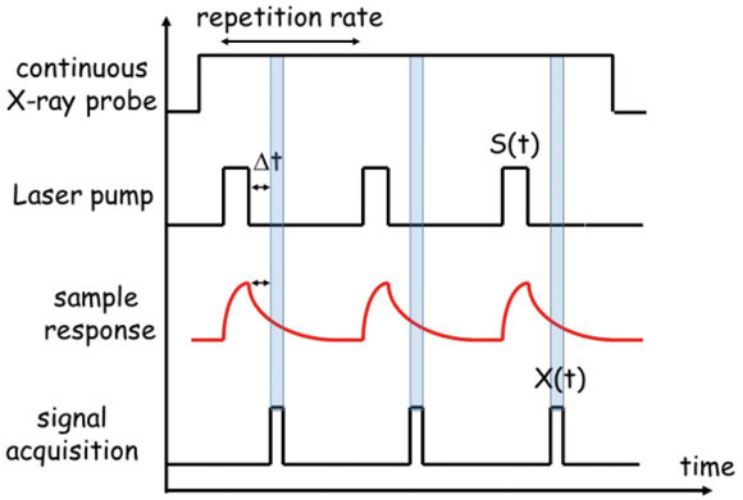


Fig. 5 A continuous X-ray beam with a scattered signal sampled by a fast-gated detector (only one delay time shown). *The figure is reproduced from Ref.* [26] *with permission from the IUCr*

2.4 Sub-picosecond and XFEL Methodologies

The interest in studying chemical and biological processes with shorter and shorter lifetimes using time-resolved crystallographic techniques continues to grow as technological advances make the picosecond and sub-picosecond time regimes accessible. The structure of the synchrotron beam itself can help with these developments. The relativistic effects on the electrons circulating the storage ring at speeds approaching that of light mean that they orbit the ring in discrete bunches [22]. Thus, the synchrotron radiation produced is naturally pulsed with a repetition rate that is determined by the period of the electron orbit around the ring. For third-generation synchrotrons, this is typically in the nanosecond to picosecond time regime. Therefore, pump-probe time-resolved experiments on species with lifetimes within this timeframe can be carried out without the need for a mechanical or electronic shutter to pulse the X-rays, and the laser repetition rate is synchronised with the repetition rate of the storage ring [40].

In order to cover all dynamic chemical processes, the ultimate aim must be to use photocrystallographic methods to investigate the initial stages of a chemical reaction that occur on femtosecond timescales [41]. The study of species with sub-picosecond lifetimes requires the development of "single-shot" diffraction methods where the whole diffraction pattern of the crystal is obtained in one X-ray pulse. Here the flux of the X-ray pulse needs to be extremely high in order to achieve a measurable pattern. Laue techniques provide sufficient flux to obtain the result, but the development of X-ray free-electron lasers (XFELs) with several orders of magnitude more flux than the most powerful synchrotrons is the obvious tool to

use to achieve this goal. The "diffract-and-destroy" approach has been developed at XFELs to study macromolecular systems [42–45]. From these and related studies, it is apparent that the direction-controlling, initial stages of a chemical reaction appear at timescales faster than are accessible through the use of synchrotrons which effectively have a 100 ps limit. The volume of data that is collected in one of these studies is enormous. For example, in the study of the initial stages of the photoexcitation of the photoactive yellow protein (PYP), which is triggered by the *trans-cis* isomerisation of the coumarin chromophore, 2.5×10^6 snapshots were recorded to 1.6 Å [46]. With these advances, and with the large number of data sets being collected accurate scaling between these data sets is required to obtain reliable results. It has been found from scaling analyses that the anisotropy of absorption following sample excitation can be pronounced and depends on the orientation of the crystals in the laser beam [47] so that this factor has to be taken into consideration if accurate analysis of molecular samples is to be achieved in the future.

3 The Beginnings of Time-Resolved Crystallography

3.1 Macromolecular Photocrystallography

The initial developments in time-resolved crystallography came in the area of macromolecular crystallography because of the interest in important biological processes. These studies required faster data collection and processing techniques than had been standard. Biological crystals are also prone to decomposition particularly as a result of X-ray radiation damage, so the use of Laue methods could achieve faster data collection with less crystal decay.

The first nanosecond time-resolved macromolecular crystallographic study using Laue techniques, with a broad range of wavelengths (white beam), was reported in 1996 when Moffat and co-workers reported a study of the photodissociation mechanism of carbon monoxide in carbon monoxy-myoglobin (MbCO) [48]. The MbCO system had previously been studied in depth by ultra-fast spectroscopic techniques, and, as a result, the photoactivity of the complex in solution had been established [49–51]. In the experiments carried out at the European Synchrotron Radiation Facility (ESRF), Moffat et al. employed a pump-probe strategy consisting of an initial 7.5 nm-wide laser pump pulse at $\lambda = 630$ nm, followed by an X-ray probe pulse timed to arrive after a specific delay (τ). This delay was varied so that six different data sets were completed at intervals of between $\tau = 4$ ns and 1.9 ms. The analysis of the data from the 4 ns and 1 μs data sets showed that regions of negative electron density appeared where the coordinated CO molecule had been. These observations confirmed the results of the earlier solution-based spectroscopic studies that photolysis of the Fe–CO bond had occurred and indicated that a similar occurs in the single crystal. The crystallographic data also showed a region of positive electron density below the heme centre, suggesting that the iron atom moves out of the heme plane as a result of the Fe–CO bond cleavage. This is consistent with the CO group

having dissociated from the iron and moved away from its binding position. The data sets collected with longer delays of above 1 μm did not show these electron density features, which is consistent with the CO recombining with the heme unit and the whole system re-relaxing back to the ground state, the whole process being completed within a few milliseconds. In a subsequent study, by the same research team, in 2001, further data sets were obtained with time delays of between 1 ns and 1 μm, and analysis of these data showed the position of transient docking sites within the heme pocket in which the photodisassociated CO sat [52]. The relative positions of these docking sites provided information on the photodissociation process and kinetic data describing the ligand recombination process.

The success of this time-resolved approach in establishing the pathways of biological processes led to further ground-breaking studies of proteins in photoactivated states with lifetimes down to hundreds of picoseconds [53, 54]. Using these pioneering techniques, it has proved possible to construct "molecular movies" that describe the full biological process in three dimensions [55].

While macromolecular time-resolved crystallography is not the focus of this chapter, the research area has continued to lead the discipline in terms of innovation taking full advantage of developments of, initially, synchrotron facilities [56–58] and, more recently, of the power of the XFELs [59–64]. The molecular crystallographers have much to learn from their macromolecular colleagues not least in the area of the treatment of crystal damage in high-intensity X-ray beams [65–67] and in the adaption of multi-crystal data collection techniques [68].

3.2 Molecular Photocrystallography

Research into photoactivated changes in single crystals of molecular compounds commenced in the 1960s when Schmidt and Cohen reported that a series of *trans*-cinnamic acid derivatives underwent irreversible [2 + 2] photodimerization cycloaddition reactions in the solid state [69–71]. In these pioneering studies, the authors highlighted the importance of the surrounding crystalline environment on the pathway of the photodimerization reaction. They presented a series of key criteria that needed to be satisfied if a single-crystal-to-single-crystal transformation was to take place without significant crystal deterioration. These criteria were generalised in the *Topochemical Postulate*, stating that the photoreaction will follow a minimum energy pathway that imparts the lowest level of steric strain to the surrounding crystal environment. This meant that only transformations that proceeded *topotactically*, that is, with the minimum amount of movement at the atomic level could occur without crystal decay. Some subsequent improvements have been made [72] and several exceptions found [73], but the *Topochemical Postulate* remains an effective guide in the design of systems that undergo high levels of photoactivated [2 + 2] cycloaddition. More recently, crystal engineering techniques have been applied to the design of systems [73] that readily undergo photoactivated [2 + 2] cycloaddition reactions in the solid state, but the basis of the *Postulate* remains

largely intact. The crystal engineering methodologies applied to the design of monomers with suitable separations and orientations have included templating methods using both metal ions and hydrogen bonds and co-crystallisations and host-guest chemistry. Because the solid-state [2 + 2] cycloaddition process is irreversible, diffraction data can be collected at various stages throughout the reaction simply by pausing the irradiation at convenient intervals. Solid-state kinetic data has been obtained in this way on cycloaddition reactions by following the photoreaction as a function of irradiation time using single-crystal X-ray diffraction methods [74–76].

Irreversible solid-state photoreactions permit the full three-dimensional structures of the starting material and of the product to be determined using conventional single-crystal X-ray crystallographic methods as long as crystal integrity is maintained throughout the process. However, for reversible dynamic processes in the solid state, effectively snapshots of the structures in their excited states must be obtained, taking into account the lifetime of the activated species. Over the last four decades, the importance of fully reversible photoactivated processes has been realised because of their application in real-world technologies [77], including sensors, read-write data storage media, non-linear optics [78], molecular switches, amphidynamic materials [77] and molecular actuators [79–81]. Examples of reversible photochemical processes include metal-metal bond-length changes [82, 83], linkage isomerisation processes [8, 23, 84, 85] and light-induced spin-state trapping behaviour [13]. The photocrystallographic studies have established that the structural changes can be promoted and controlled in the solid state, with the concomitant control over physical properties such as colour [86, 87], luminescence [8, 17, 24, 86, 87] and refractive index [88–90].

Steady-state and pseudo-steady-state photocrystallographic techniques have been used to identify metastable species with much of the research focussing on the identification of metastable linkage isomers and on the products of light-induced excited spin-state trapping (LIESST) experiments. This work is covered in previous chapter of this book by Skelton et al. and will not be discussed in detail here other than to give a brief outline of the overall findings that can be related to the faster time-resolved photocrystallographic experiments that will be described in the next section.

The majority of photocrystallographic studies of transition metal complexes that undergo linkage isomerism under photoactivation have focussed on nitrosyl [91–94], sulphur dioxide [17, 85, 95] and nitro complexes [8]. Some general conclusions as to the processes involved can be drawn from these studies. In all cases the percentage of conversion obtained and the isomer formed are highly dependent on the wavelength of light used and the temperature at which the experiment is carried out. For a single-crystal-to-single-crystal process to occur, there is no change in the crystal system and the unit cell parameters do not change by more than 2%. There is a temperature at which the metastable limit is reached above which the excited state has a finite lifetime before returning to the ground state. The interconversion is also dependent on the flexibility of the crystal lattice and on the steric and electronic environment of the ambidentate ligand and the metal centre, as evidenced by only small changes in cell volume being observed. There needs to be sufficient space in

the lattice for the ligand to switch between one form and the other, and intermolecular interactions should either not be particularly strong or be flexible perhaps through a change in temperature [88, 96]. The accessible volume within the lattice and the flexibility of the lattice under experimental conditions relates back to the concept of the *reaction cavity* first proposed by Cohen [72] and then exemplified by Ohashi who showed subsequently that the *reaction cavity* could flex during the course of the reaction and that changes in temperature could significantly affect the process [97]. For example, lowering the temperature would cause lattice contraction, reducing the cavity size, and might thus "switch off" the reaction.

Light-induced excited spin-state trapping (LIESST) studies map the low spin-high spin interconversions in transition metal complexes with d^4-d^7 electronic configurations. The spin crossover phenomenon, with the associated change in magnetic properties, is usually activated thermally with the high spin (HS) to low spin (LS) interconversion occurring on cooling the complex to below some critical temperature. A significant number of these complexes can then undergo a photoactivated conversion from the LS state to a metastable HS state, via a long-lived triplet excited state. The photoinduced phenomenon was first identified in a study of $[Fe(1\text{-ptopyltetrazole})_6][BF_4]_2$ using Mössbauer spectroscopy [98]. The phenomenon is observable in both solution and in the solid state.

Single-crystal crystallographic and powder diffraction studies of the LIESST phenomenon soon followed [99–105], taking advantage of cryoscopic advances that allowed crystals to be relatively easily cooled below liquid nitrogen temperatures; many LIESST transitions occurring below 80 K. Since the early 2000s, the topic has continued to develop with a range of techniques being used to analyse the LIESST phenomenon [13, 106]. Of additional interest are LIESST complexes that display reverse switching from their photoinduced HS state back to a low temperature LS state which is induced via both temperature changes and "reverse-LIESST" processes, involving further irradiation of the excited LIESST state with a different excitation wavelength [107]. This two-way switching using different wavelengths of light has potential applications, making these metastable state species potential candidates for photoswitchable molecular devices.

3.3 Time-Resolved Molecular Photocrystallographic Studies

For successful pump-probe molecular photocrystallographic studies, a material in which the molecules can undergo a fast, fully reversible switching process is required, and there should only be a small change in unit cell dimensions during the process. The robustness of the crystal under light and X-ray radiation is of primary importance if sufficient data is to be obtained and the excited state structure solved to atomic resolution.

3.3.1 Studies with Monochromatic X-Ray Radiation

Pump-probe X-ray experiments on molecular species, using monochromatic radiation rather than Laue techniques, were first described in 2001. Techert et al. reported a time-resolved study of transient changes in N,N-dimethylaminonitrile (DMABN) using stroboscopic X-ray powder diffraction (XRPD) methods [108]. Photoactivation of DMABN at $\lambda = 267$ nm resulted in an ultra-fast structural change as a result of the electronic excitation, and relaxation back to the ground state occurs on the picosecond timescale. The relaxation process was structurally analysed using synchrotron radiation with the beam being chopped mechanically and synchronised with light pulses from a Ti:sapphire laser, with pump-probe delay times ranging from −240 to 1,500 ps, allowing for several data points to be collected. The reflections that showed the greatest intensity changes during the structural rearrangement were identified, and the authors followed these intensity changes with increasing delay times. The data showed a conversion level to the excited state of between 28 and 32% at short time delays, which is consistent with spectroscopic data. Refinement of the powder data using the Rietveld method showed that the relaxation occurred via a change in the inversion angle of the molecule and a rotation of the methyl group (as shown in Fig. 6), with a total structural relaxation time of 520 ps.

Coppens was among the first to take advantage of the higher intensity of synchrotron X-ray radiation and in 2002 reported the first single-crystal X-ray diffraction study of a species with a microsecond lifetime. The results of pump-probe experiments on salts of the $[Pt_2(pop)_4]^{4-}$ anion (pop = $[H_2P_2O_5]^{2-}$) showed that structural distortions in the anion are induced by photoactivation with 355 nm light, producing a triplet excited state with a microsecond lifetime [83]. The key change in the structure upon excitation was the shortening of the Pt–Pt bond which has previously been proposed from electronic [109] and Raman spectroscopy [110] and XAFS [111] studies, the latter suggesting a reduction in Pt–Pt bond length of 0.52 Å. The pump-probe time-resolved single-crystal X-ray study was carried out on the tetra-tetraethylammonium salt of $[Pt_2(pop)_4]^{4-}$ at a temperature of 17 K and 33 μs wide light pulses from a Nd/YAG laser, at a repetition rate of 5,100 Hz, and using the "*light-on/light-off*" data collection strategy. Analysis of the excited state data showed a 2% level of excitation under the experimental conditions. Upon refinement of the Pt positions, the rest of the structure being treated as a rigid group, the Pt–Pt distance was found to shorten by 0.28(9) Å with a concomitant rotation of 3° about the Pt–Pt vector. The Pt–Pt distance has subsequently been confirmed by a second diffraction experiment [112] on the related salt $(n\text{-}Bu_4N)_2H_2[Pt_2(pop)_4]$ where the Pt–Pt reduction in length was recorded at 0.23 Å. More recently, further confirmation has come from scattering measurements of an aqueous solution [113] and a time-resolved EXAFS study [114] that showed the reduction in Pt–Pt distance from that found in the ground state structure were 0.24 Å and 0.31 Å, respectively.

Another substantial reduction in metal-metal bond length was reported for the Rh–Rh bond length in $[Rh_2(dimen)_4][PF_6]_2$.MeCN (dimen = 1,8-diisocyanomethane)

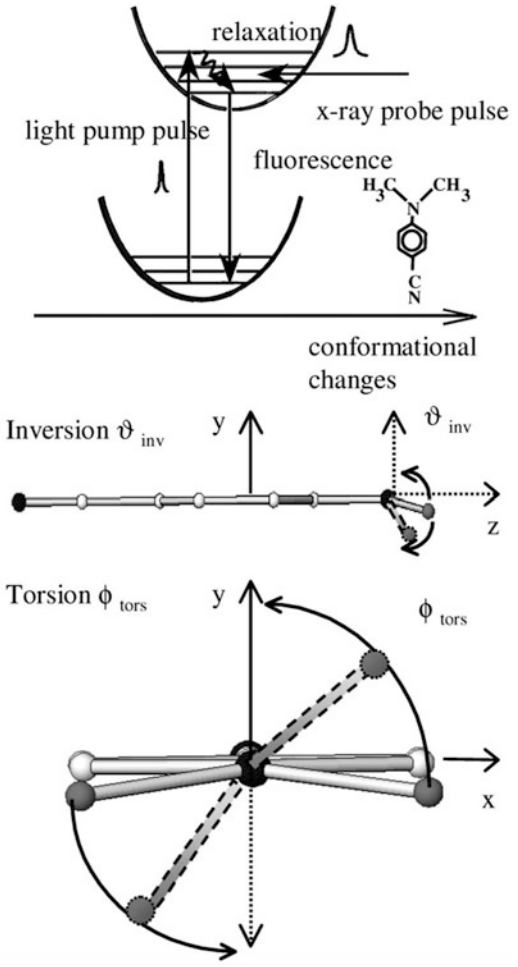


Fig. 6 Top: The principle of TR X-ray diffraction on the excited state PES of DMABN crystals. Bottom: the intramolecular degrees of freedom, which contribute to the relaxation process (inversion qinv and torsion ftors). C atoms of the phenyl moiety are given as open, N atoms as black and the amino C as grey circles. Note that the H atoms are not shown, since they do not contribute to the X-ray diffraction signal. *Reproduced from Ref.* [108] *with permission from the ACS*

upon photoactivation at 23 K using 335 nm laser pulses. The maximum level of excitation reached was 2.5%, and the Rh–Rh bond length reduction was a remarkable 0.86 Å and a bond rotation of ca. 13° [82]. Previous spectroscopic studies on the cation had indicated a lifetime of ca. 11 μs for the triplet excited state species [115] which is in excellent agreement with the experimental value of 11.7 μs at 23 K. DFT calculations conducted at the same time as the photocrystallography experiment proposed a slightly greater decrease in the Rh–Rh distance upon excitation as illustrated in Fig. 7. The difference between the experimental and theoretical results may indicate that the steric effects from the crystalline environment may modify the structural response to the excitation process.

Further photocrystallographic studies employing stroboscopic methods were reported by Coppens et al. including an investigation of photoinduced structural

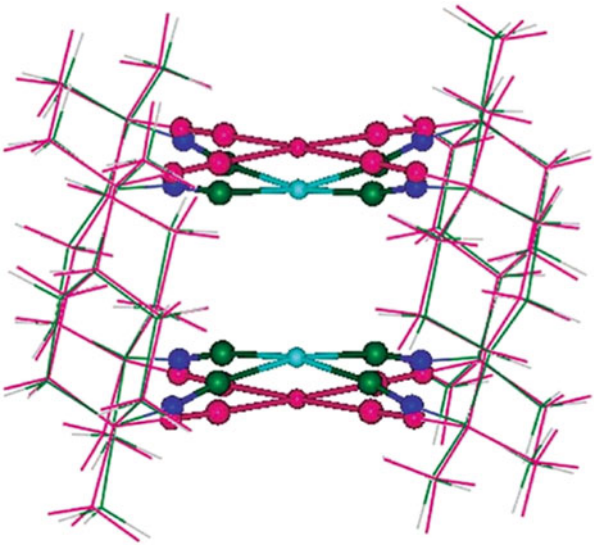


Fig. 7 Theoretical ground state (in pink) and excited state (green and blue colours) geometries of the $[Rh_2(1,8\text{-diisocyanomenthane})_4]^{2+}$ cation. *Reproduced from Ref.* [82] *with permission from the Royal Society of Chemistry*

changes in the trimeric complex $[Cu_3(3,5\text{-}(CF_3)_2\text{pyrazolate})_3]$ [116]. This system differs from the examples discussed above because the photoexcitation occurs intermolecularly between two neighbouring copper trimer molecules, as opposed to the intramolecular processes in the di-nuclear species. Photoactivation at $\lambda = 355$ nm and 17 K promotes the production of excited state species with microsecond lifetimes, involving a rearrangement of the trimer molecules into pairs such that one inter-planar Cu...Cu distance is reduced by 0.65 Å while the next Cu...Cu contact lengthens by ca. 0.30 Å. In a separate study, conducted in 2009, significant structural distortions in the complex $[Cu(dmp)(dppe)](PF_6)$, (dmp = 2,9-dimethyl-1,10-phenanthroline) are described [117]. The complex crystallises with two, crystallographically independent molecules in the asymmetric unit, whose structural response on photoactivation is interestingly different. These differences have been attributed to different constraining effects from the surrounding crystalline environment for each of the independent molecules. In general, upon irradiation the Cu cation is observed to "flatten out", and a concomitant increase in the average Cu–P bond length is observed by comparison of the diffraction data from the ground and excited states (Fig. 8). The changes are expected to be the result of charge-transfer between the dmp and dppe ligands and were determined from an excited state population of ca. 7–10% in the single crystal.

These studies highlighted the importance of the crystalline environment on the solid-state photoactivation process and confirmed the significance of the "reaction cavity" hypothesis discussed earlier [97, 118]. In order to exploit this aspect and to use it to increase the level of conversion to the excited state species, photoactive

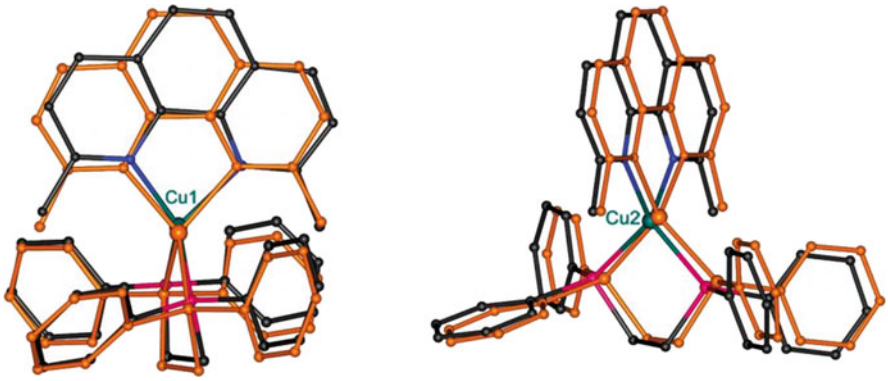


Fig. 8 Excited state geometries of the two independent molecules (orange) superimposed on the ground state of the complex (Cu, green; C, black; P, purple; N, blue). Slightly different views are shown to illustrate the change in rocking distortion (left) and the displacement of the phenanthroline ligand from its ground state plane (molecule 2, right) upon excitation. *Reproduced from Ref.* [117] *with permission from the American Chemical Society*

complexes have been embedded as guests into arrays and framework materials. It is of importance that the host frameworks should be inert to photoactivation but effectively dilute the concentration of the photoactive species in the material. The host-guest arrangement has significant advantages for promoting the retention of crystal integrity during the process and provides greater freedom for the guest to rearrange without resulting in steric clashes with adjacent molecules in the pure material. Additionally, the dilution of the photoactive species reduces the number of photons that are required to maximise excitation, leading to more efficient photoactivation and, hopefully, increased conversion percentages [119]. It should also be noted that this approach has the effect of isolating the photoactive molecules from one another, producing quite a different environment to that experienced in crystals of the guest molecule, so differences in the physical properties of the pure compound and of the host-guest complex should be expected. Coppens et al. have investigated a number of species by this approach, via both static and dynamic photocrystallographic techniques [94, 119–121].

There have been a number of other studies using "crystal engineering" techniques involving the use of molecular cages and flasks [122] and metal organic frameworks (MOFs) to trap transient and highly reactive species [123]. Kawano et al. used synchrotron X-ray radiation to identify the coordinatively unsaturated "η^5-$(C_5H_4Me)Mn(CO)_2$" moiety in a designed self-assembled coordination cage [124]. Photoirradiation of a [η^5-$(C_5H_4Me)Mn(CO)_3$] guest molecule, within a single crystal, at 100 K, using 365 nm light, resulted in the dissociation of a carbonyl ligand remarkably without loss of crystallinity, and peaks in the electron density difference map could be attributed to free carbon monoxide. The crystallographic results were supported by a solid-state IR study. The same group has subsequently gone on to identify an unstable imine [125] and a transient hemiaminal [126], both trapped in pre-designed porous networks. They have also demonstrated the suppression of

rapid *cis-trans* isomerism in the dimeric coordination complex $[(\eta^5\text{-indenyl})Ru(CO)_2]_2$ when this molecule is trapped in a self-assembled coordination cage [127] and the conversion of an overcrowded chromic alkene to a metastable twisted conformation, upon photoactivation, when incorporated into a tetrahedrally symmetric coordination cage [128]. Champness et al. have also exploited the power of synchrotron-based X-ray crystallography to support time-resolved IR studies to probe the *fac-mer* isomerism of $[M(diimine)(CO)_3X]$ (M = Re or Mn; X = Cl or Br) units immobilised in a MOF and shown the presence of the *mer*-isomer in the photoactivated crystalline solid [129]. Subsequently, they have shown that the coordination polymer $[(Cu(DMF)(H_2O)][LRe(CO)_3Cl].DMF]_\infty$ (L = 2,2′-bipyridine-5,5′-dicarboxylic acid) undergoes an irreversible photoinduced charge transfer process. Time-resolved IR spectroscopy was used to identify the nature of this photoinduced process and how, under suitable conditions, it is possible to initiate irreversible modification of the crystal through induction of the charge transfer process. By using the photoinduced process, which arises purely as a result of the structure of the coordination network, it was possible to write on the surfaces of crystals [130].

3.3.2 Studies Using Laue Diffraction Techniques

Among the first time-resolved photocrystallographic studies using Laue diffraction techniques, designed to probe shorter timescales and reduce crystal damage, was carried out on a di-rhodium complex, $[Rh_2(\mu\text{-PNP})_2(PNP)_2][BPh_4]$ ($PNP = CH_3N(P(OCH_3)_2)_2$), which was studied photocrystallographically on the 100 ps timescale [29] using Laue diffraction methods. The crystal was pumped with 35 ps pulses of a Ti:sapphire laser tuned to a wavelength of 337 nm, at a temperature of 225 K, followed by a single 100 ps-wide X-ray pulse after a 100 ps delay. The lifetime of the excited state species was ca. 35 μs. In this case, the experimental results showed a shortening of the Rh–Rh distance of 0.136(8) Å upon excitation, and the results are quantitatively supported by quantum-mechanical calculations. The study shows similar, but smaller, trends to those described for $[Rh_2(dimen)_4][PF_6]_2$.MeCN above [82]. The deconvolution of the overlapping reflections was achieved during the data reduction, and the structure was refined using the RATIO method [36].

A Cu(I) complex, $[Cu(1,10\text{-phenanthroline})(PPh_3)_2][BPh_4]$, was also studied using Laue diffraction methods [131] and compared to the $[Cu(dmp)(dppe)]^+$ cation that had been studied previously using monochromatic methods [117]. This phenanthroline complex also crystallises with two independent cations in the asymmetric unit, with one molecule in a more sterically constrained environment. Photoactivation in the single crystal was expected to induce a MLCT transition, resulting in transient structural changes. The two independent molecules are again observed to undergo different structural responses upon activation with $\lambda = 390$ nm light at 90 K. Data were collected using the single-pulse Laue method, the data was analysed [132], and the results showed considerable distortion in the less restricted Cu molecule, while no significant changes were observed in the cation adopting the second, more confined arrangement (Fig. 9). Theoretical studies confirmed that the

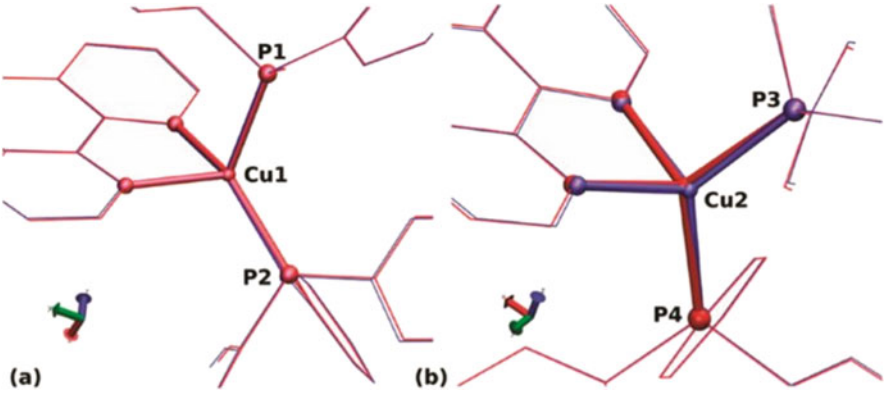


Fig. 9 Comparison of the ground state (in blue) and the excited state (in red) structures of the two independent molecules at 90 K. (**a**) Molecule A; (**b**) molecule B. *Reproduced from Ref.* [131] *with permission from the American Chemical Society*

different responses were to be expected resulting from the interaction of the photoactive species with the surrounding lattice, with far greater distortions predicted in the free molecule than those observed experimentally in the solid state. For both this and the study on $[Rh_2(\mu\text{-}PNP)_2(PNP)_2][BPh_4]$, the experimental standard deviations for data collected by the Laue method are much improved compared to those seen with monochromatic time-resolved techniques. These improvements highlight the greater accuracy of time-resolved diffraction data collected by Laue methods for short lifetime species, which is mainly attributed to reduced levels of laser heating when adopting the single-shot approach [133].

The use of pink-beam Laue diffraction has also been applied to the time-resolved photocrystallographic study of mixed metal polynuclear complexes. The triplet excited state of the tetranuclear d^{10}-d^{10} complex $Ag_2Cu_2L_4$ (L = 2-diphenylphosphino-3-methylindole) has been investigated with a Laue pump-probe technique with an 80 ps time resolution at 90 K [134]. The lifetime of 1 μs is accompanied by significant changes in the metal framework, with an Ag…Cu distance shortening by 0.59(3) Å, which suggests an increase in the argentophilic interactions (Fig. 10). The photocrystallographic study was accompanied by theoretical calculations which confirm that the strengthening of the Ag…Ag interaction is caused by ligand-to-metal charge transfer (LMCT).

Most recently, the luminescent properties of a tetranuclear Cu(I) benzoate complex have been investigated by a combination of time-resolved spectroscopy and crystallography. The complex $[Cu_4(PhCO_2)_4]$ displays luminescent thermochromism, with red phosphorescence at room temperature that turns green on lowering the temperature to 90 K [135]. The low-energy triplet state has been assigned to a cluster-centred triplet state, and the emission from this state matches the experimental red band observed at 660–715 nm. The computed next highest triplet excited state occurs close to the experimental value at 545 nm. The two excited states exhibit MLCT and LMCT characteristics, particularly in their solid-

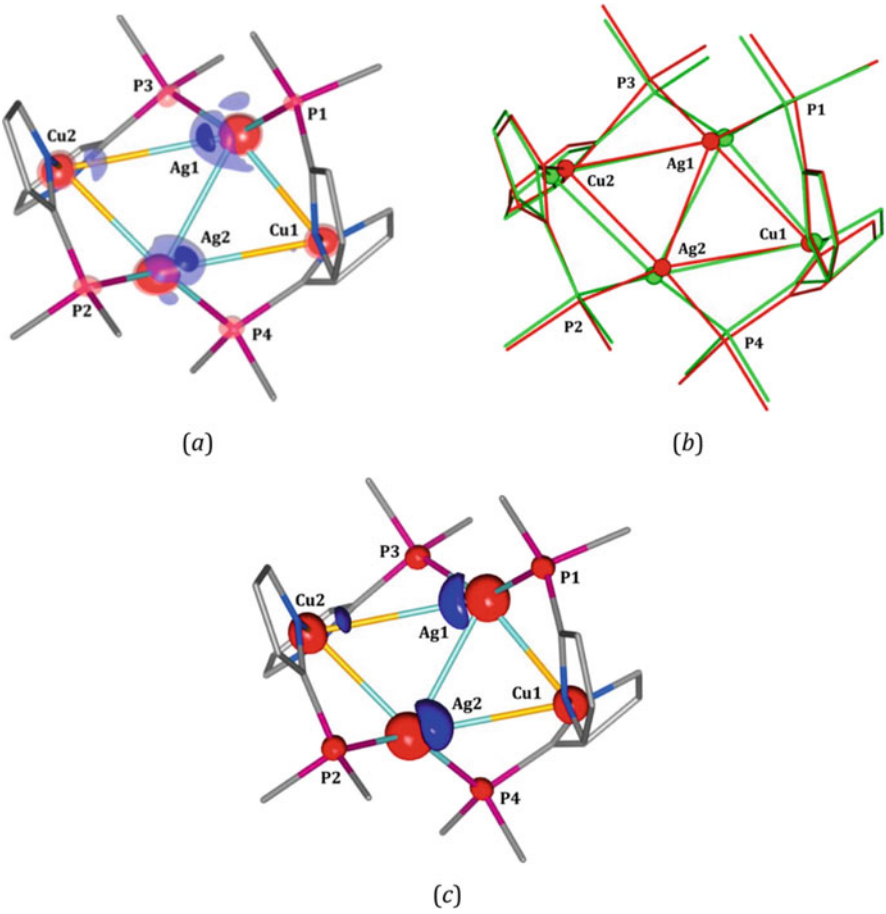


Fig. 10 (**a**) Photodifference map ($F_o^{ON}-F_o^{OFF}$) of the complex showing atomic shifts upon excitation (solid isosurfaces, ±0.55 e·Å^{-3}; semi-transparent, ±0.35 e·Å^{-3}; blue, positive; red, negative). (**b**) Refined excited state geometry related to that of the ground state crystal structure (green, ground state; red, excited state; methylindole ligands were omitted for clarity). (**c**) Photodeformation map ($F_c^{ON}-F_c^{OFF}$) based on the refined model parameters (isosurfaces, ± 0.30 e·Å^{-3}; blue, positive; red, negative; $k_B = 1.06$). *Reproduced from Ref.* [134] *with permission from the American Chemical Society*

state geometries, as evidenced from the Laue photocrystallographic study, at 90 and 225 K, with 355 and 360 nm light irradiation, which shows the expected Cu. . .Cu contraction. Again, there are two independent molecules in the crystallographic asymmetric unit, and they show slightly different distortions because of their different crystalline environments.

With recent technological advances, it is not just very fast timescale photocrystallographic molecular transformations that have been carried out. It has become possible to study the dynamics of LIESST [136], and the results show

significant differences to results obtained from those of conventional diffraction studies of long-lived photoinduced high spin states. Ultra-fast spin-state photoswitching in two crystalline polymorphs of the octahedral Fe(III) complex [(TPA)Fe(TCC)] (TPA = tris(2-pyridylmethyl)amine and TCC = 3,4,5,6-tetrachloroactecholate dianion) (Fig. 11) has been studied by femtosecond optical spectroscopy and picosecond X-ray diffraction [39]. In these experiments the X-ray pulses were generated by a fast chopper, and data was collected with a series of different delay times between the laser pump and the X-ray probe. The time

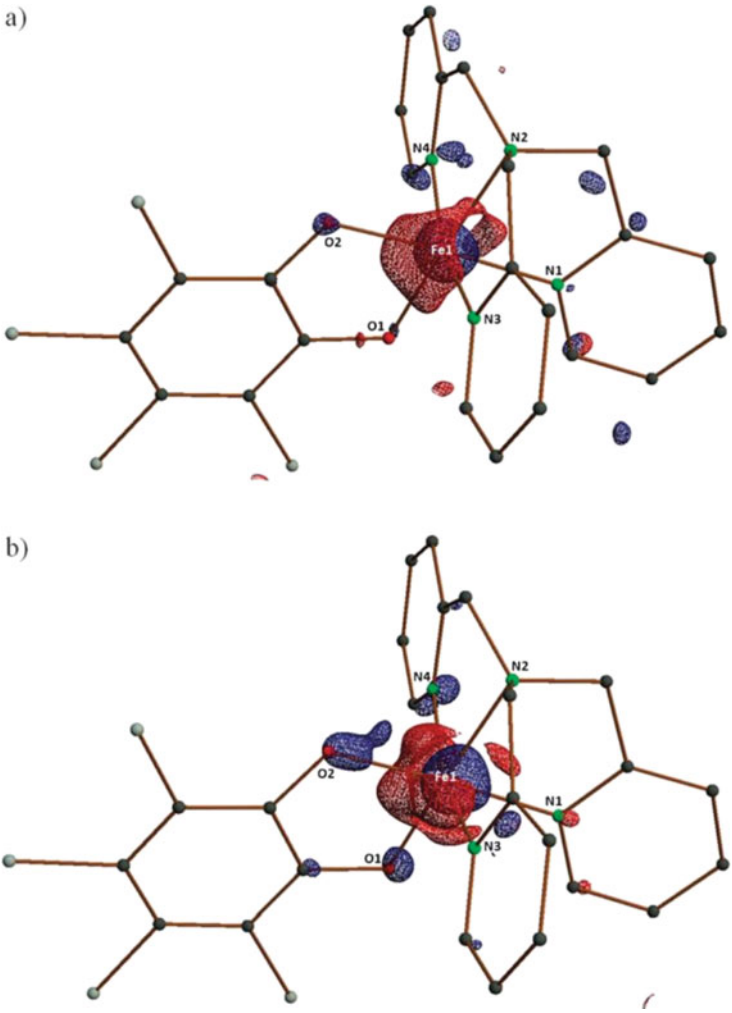


Fig. 11 Photodifference maps obtained for the monoclinic polymorph with isosurfaces (red positive, blue negative) of (**a**) ±0.14 eÅ^3 for the 500 ps data and (**b**) ±0.46 eÅ^3 for the 50 μs data. *Reproduced from Ref.* [39] *with permission from the Royal Society of Chemistry*

dependence of the lattice parameters was measured using partial data collected for each delay time with 60 frames with 10 s of exposure for every 1° step of the diffractometer ϕ axis. Typical excitation densities were 150 μm^{-2}, with a laser diameter of ca. 500 μm (FWHM). After data processing of the full data sets for the monoclinic polymorph, photodifference maps were calculated for the 500 ps and 50 μs time delays. The maps show significant changes in electron density upon excitation (Fig. 11). The response to the excitation at 500 ps is shown by the sideways shift of electron density from the Fe1 atom towards O2 (Fig. 11a). This shift becomes even more pronounced at the 50 μs time point (Fig. 11b). There is also an increase in the Fe–N distances by ca. 0.05 Å. The estimated population of the excited state at 500 ps is in the range 1.5–2%, and this increases to ca. 10.5% after 50 μs. Overall, the results show significant shifts of the Fe atom and the two O atoms upon excitation, and the results are consistent with the complementary spectroscopy that was undertaken.

In a subsequent study of the Fe(III) spin-crossover material, [Fe(3-MeO-SalEen)$_2$][PF_6], the switch from LS to HS only occurs at the molecular level as clearly shown by the linear dependence of the fraction of photoswitched molecules with the excitation density as well as with the initial fraction of low spin molecules. The inter-system crossing from the photoexcited LS (S = 1/2) to HS (S = 5/2) occurs within approximately 200 fs and is accompanied by coherent non-equilibrium vibrational relaxation in the photoinduced HS state. These results reveal similar dynamical features to those already reported for LIESST in related Fe(II) systems [137]. The activation of coherent molecular vibrations is an essential requirement for reaching the HS potential on the timescale of molecular motions, whereas their fast damping allows an efficient trapping in the HS potential [138]. The observed coherent oscillations are attributed to photoinduced molecules in the HS states, and the results are supported by Raman spectroscopy at thermal equilibrium and DFT analyses of molecular vibrations and TD-DFT calculations of optical absorption.

4 Conclusions

Time-resolved crystallography has developed extensively over the last three decades, and the results presented in this chapter show that under favourable conditions it is possible to obtain full three-dimensional structural data on chemical species that have lifetimes of microseconds or less [139]. These advances, coupled with similar advances in instrumentation and computer power, open up possibilities for monitoring chemical processes in the solid state in a way that has not previously been possible. At present the restrictions of maintaining crystal integrity remain, and the reactions and processes are mostly limited to the interaction of the solid to external media such as light, as discussed in this chapter, but also to changes in temperature and pressure and the influence of magnetic and electric fields (although crystal integrity is retained in some solid-gas reactions). The potential is enormous,

and the topic is ripe for development. As has already been indicated, the new frontier in photocrystallography over the next several decades is the use of X-ray free-electron laser sources to carry out serial femtosecond studies [41]. The use of XFELs to track chemical reactions by serial femtosecond pump-probe X-ray crystallography by rapidly recording diffraction patterns at closely spaced time intervals has already proved successful in macromolecular crystallography [42, 43, 45]. Once the scaling issues related to using multiple crystals in serial processes are resolved and diffraction patterns better than atomic resolution [47] are obtained, the prospect for carrying out molecular crystallography at an XFEL is extremely promising.

In the meantime, for the slower timescale processes (milliseconds to minutes), detector development will facilitate the use of laboratory sources [26, 140], rather than the need for synchrotrons. This will mean that many more "routine" time-resolved experiments can be carried out and there are thousands of systems which would benefit from structural dynamic studies. Additionally, the monitoring of dynamic processes using multiple techniques simultaneously, such as the combination of photocrystallography with emission spectroscopy or Raman spectroscopy, will allow the detailed processes within the molecule as well as changes in the crystalline environment to be monitored in exquisite detail.

Acknowledgements PRR gratefully acknowledges the support of the Engineering and Physical Sciences Research Council (EP/K004956) and the Diamond Light Source for the provision of beamtime.

Dedication This chapter is dedicated to Professor Alan J. Welch of the Chemistry Department, Heriot-Watt University, Edinburgh, Scotland, on his retirement. Alan was completing his Ph.D. in Chemical Crystallography in Mike Hursthouse's research group at Queen Mary College, London, when I joined the group to start mine. Alan guided me through the first year of my Ph.D., and I am very grateful to him for passing on all his crystallographic knowledge. We have kept in touch ever since, and it is nice to have this opportunity to recognise Alan's outstanding contribution to structural chemistry.

References

1. Burgi HB, Dunitz JD, Shefter E (1973) Geometrical reaction coordinates. II Nucleophilic addition to a carbonyl group. J Am Chem Soc 95(15):5065–5067
2. Clegg W, Blake AJ, Cole JM, Evans JSO, Main P, Parsons S, Watkin DJ (2009) Crystal structure analysis: principles and practice. Oxford University Press, Oxford. Incorporated
3. Weller MT, Mark TW (2017) Young NA, Nigel AY (eds) Characterisation methods in inorganic chemistry. Oxford University Press, Oxford
4. Naumov P, Bharadwaj PK (2015) Single-crystal-to-single-crystal transformations. CrstEngComm 17(46):8775–8775
5. Chaudhary A, Mohammad A, Mobin SM (2017) Recent advances in single-crystal-to-single-crystal transformation at the discrete molecular level. Cryst Growth Des 17(5):2893–2910
6. Bryant MJ, Skelton JM, Hatcher LE, Stubbs C, Madrid E, Pallipurath AR, Thomas LH, Woodall CH, Christensen J, Fuertes S, Robinson TP, Beavers CM, Teat SJ, Warren MR, Pradaux-Caggiano F, Walsh A, Marken F, Carbery DR, Parker SC, McKeown NB, Malpass-

Evans R, Carta M, Raithby PR (2017) A rapidly-reversible absorptive and emissive vapochromic Pt(II) pincer-based chemical sensor. Nat Commun 8
7. Lee JH, Park S, Jeoung S, Moon HR (2017) Single-crystal-to-single-crystal transformation of a coordination polymer from 2D to 3D by [2 + 2] photodimerization assisted by a coexisting flexible ligand. CrstEngComm 19(27):3719–3722
8. Hatcher LE, Skelton JM, Warren MR, Raithby PR (2019) Photocrystallographic studies on transition metal nitrito metastable linkage isomers: manipulating the metastable state. Acc Chem Res 52(4):1079–1088
9. Commins P, Desta IT, Karothu DP, Panda MK, Naumov P (2016) Crystals on the move: mechanical effects in dynamic solids. Chem Commun 52(97):13941–13954
10. Moggach SA, Parsons S (2009) High pressure crystallography of inorganic and organometallic complexes. Spectroscopic properties of inorganic and organometallic compounds. The Royal Society of Chemistry 40:324–354
11. Woodall CH, Beavers CM, Christensen J, Hatcher LE, Intissar M, Parlett A, Teat SJ, Reber C, Raithby PR (2013) Hingeless negative linear compression in the mechanochromic gold complex [(C6F5Au)2(μ-1,4-diisocyanobenzene)]. Angew Chem 125(37):9873–9876
12. Shepherd HJ, Rosa P, Vendier L, Casati N, Létard J-F, Bousseksou A, Guionneau P, Molnár G (2012) High-pressure spin-crossover in a dinuclear Fe(II) complex. Phys Chem Chem Phys 14 (15):5265–5271
13. Halcrow MA (2011) Structure:function relationships in molecular spin-crossover complexes. Chem Soc Rev 40(7):4119–4142
14. Legrand Y-M, van der Lee A, Masquelez N, Rabu P, Barboiu M (2007) Temperature induced single-crystal-to-single-crystal transformations and structure directed effects on magnetic properties. Inorg Chem 46(22):9083–9089
15. Reinoso S, Artetxe B, Gutierrez-Zorrilla JM (2018) Single-crystal-to-single-crystal transformations triggered by dehydration in polyoxometalate-based compounds. Acta Crystallogr C 74(11):1222–1242
16. Coppens P, Fomitchev DV, Carducci MD, Culp K (1998) Crystallography of molecular excited states. Transition-metal nitrosyl complexes and the study of transient species. J Chem Soc Dalton 6:865–872
17. Cole JM, Irie M (2016) Solid-state photochemistry. CrstEngComm 18(38):7175–7179
18. Hatcher LE, Raithby PR (2014) Dynamic single-crystal diffraction studies using synchrotron radiation. Coord Chem Rev 277:69–79
19. Cole JM (2008) Photocrystallography. Acta Crystallogr A 64:259–271
20. Coppens P (2009) The new photocrystallography. Angew Chem Int Ed Engl 48 (24):4280–4281
21. Naumov P, Sahoo SC, Zakharov BA, Boldyreva EV (2013) Dynamic single crystals: kinematic analysis of photoinduced crystal jumping (the photosalient effect). Angew Chem Int Ed 52(38):9990–9995
22. Clegg W (2000) Synchrotron chemical crystallography. J Chem Soc Dalton 19:3223–3232
23. Hatcher LE, Raithby PR (2013) Solid-state photochemistry of molecular photo-switchable species: the role of photocrystallographic techniques. Acta Crystallogr C 69:1448–1456
24. Hatcher LE, Skelton JM, Warren MR, Stubbs C, da Silva EL, Raithby PR (2018) Monitoring photo-induced population dynamics in metastable linkage isomer crystals: a crystallographic kinetic study of [Pd(Bu4dien)NO2]BPh4. Phys Chem Chem Phys 20(8):5874–5886
25. Fullagar WK, Wu G, Kim C, Ribaud L, Sagerman G, Coppens P (2000) Instrumentation for photocrystallographic experiments of transient species. J Synchrotron Radiat 7:229–235
26. Casaretto N, Schaniel D, Alle P, Wenger E, Parois P, Fournier B, Bendeif E, Palin C, Pillet S (2017) In-house time-resolved photocrystallography on the millisecond timescale using a gated X-ray hybrid pixel area detector. Acta Crystallogr B Struct Sci Cryst Eng Mater 73:696–707
27. Husheer SLG, Cole JM, d'Almeida T, Teat SJ (2010) A prototype chopper for synchrotron time-resolved crystallographic measurements. Rev Sci Instrum 81(4)

28. Yorke BA, Beddard GS, Owen RL, Pearson AR (2014) Time-resolved crystallography using the Hadamard transform. Nat Methods 11(11):1131–1134
29. Makal A, Trzop E, Sokolow J, Kalinowski J, Benedict J, Coppens P (2011) The development of Laue techniques for single-pulse diffraction of chemical complexes: time-resolved Laue diffraction on a binuclear rhodium metal-organic complex. Acta Crystallogr A 67:319–326
30. Schmokel MS, Kaminski R, Benedict JB, Coppens P (2010) Data scaling and temperature calibration in time-resolved photocrystallographic experiments. Acta Crystallogr A 66:632–636
31. Coppens P, Fournier B (2015) New methods in time-resolved Laue pump-probe crystallography at synchrotron sources. J Synchrotron Radiat 22:280–287
32. Bourgeois D, Wagner U, Wulff M (2000) Towards automated Laue data processing: application to the choice of optimal X-ray spectrum. Acta Crystallogr D Struct Biol 56:973–985
33. Coppens P, Vorontsov II, Graber T, Gembicky M, Kovalevsky AY (2005) The structure of short-lived excited states of molecular complexes by time-resolved X-ray diffraction. Acta Crystallogr A 61:162–172
34. Benedict JB, Makal A, Sokolow JD, Trzop E, Scheins S, Henning R, Graber T, Coppens P (2011) Time-resolved Laue diffraction of excited species at atomic resolution: 100 ps single-pulse diffraction of the excited state of the organometallic complex Rh-2(mu-PNP)(2)(PNP)(2) center dot BPh4. Chem Commun 47(6):1704–1706
35. Coppens P, Pitak M, Gembicky M, Messerschmidt M, Scheins S, Benedict J, Adachi S, Sato T, Nozawa S, Ichiyanagi K, Chollet M, Koshihara S (2009) The RATIO method for time-resolved Laue crystallography. J Synchrotron Radiat 16:226–230
36. Vorontsov I, Pillet S, Kaminski R, Schmokel MS, Coppens P (2010) LASER – a program for response-ratio refinement of time-resolved diffraction data. J Appl Cryst 43:1129–1130
37. Fournier B, Sokolow J, Coppens P (2016) Analysis of multicrystal pump-probe data sets. II Scaling of ratio data sets. Acta Crystallogr A Found Adv 72:250–260
38. Fournier B, Coppens P (2014) On the assessment of time-resolved diffraction results. Acta Crystallogr A Found Adv 70:291–299
39. Collet E, Moisan N, Balde C, Bertoni R, Trzop E, Laulhe C, Lorenc M, Servol M, Cailleau H, Tissot A, Boillot ML, Graber T, Henning R, Coppens P, Buron-Le Cointe M (2012) Ultrafast spin-state photoswitching in a crystal and slower consecutive processes investigated by femtosecond optical spectroscopy and picosecond X-ray diffraction. Phys Chem Chem Phys 14(18):6192–6199
40. Coppens P, Iversen B, Larsen FK (2005) The use of synchrotron radiation in X-ray charge density analysis of coordination complexes. Coord Chem Rev 249(1–2):179–195
41. Coppens P (2017) The dramatic development of X-ray photocrystallography over the past six decades. Struct Dyn 4(3)
42. Spence JCH (2017) X-ray lasers in biology: structure and dynamics. In: Hawkes PW (ed) Advances in imaging and electron physics, vol 200. Academic Press Inc, pp 103–152
43. Spence JCH (2017) XFELs for structure and dynamics in biology. Iucrj 4:322–339
44. Pande K, Hutchison CDM, Groenhof G, Aquila A, Robinson JS, Tenboer J, Basu S, Boutet S, DePonte DP, Liang MN, White TA, Zatsepin NA, Yefanov O, Morozov D, Oberthuer D, Gati C, Subramanian G, James D, Zhao Y, Koralek J, Brayshaw J, Kupitz C, Conrad C, Roy-Chowdhury S, Coe JD, Metz M, Xavier PL, Grant TD, Koglin JE, Ketawala G, Fromme R, Srajer V, Henning R, Spence JCH, Ourmazd A, Schwander P, Weierstall U, Frank M, Fromme P, Barty A, Chapman HN, Moffat K, van Thor JJ, Schmidt M (2016) Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. Science 352(6286):725–729
45. Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia F, 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 MN, 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 XY, Grotjohann I, Holton JM, Barends TRM, 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 JCH (2011) Femtosecond X-ray protein nanocrystallography. Nature 470(7332):73–U81

46. Schmidt M, Pande K, Basu S, Tenboer J (2015) Room temperature structures beyond 1.5 angstrom by serial femtosecond crystallography. Struct Dyn 2(4)
47. Coppens P, Fournier B (2015) On the scaling of multicrystal data sets collected at high-intensity X-ray and electron sources. Struct Dyn 2(6)
48. Srajer V, Teng TY, Ursby T, Pradervand C, Ren Z, Adachi S, Schildkamp W, Bourgeois D, Wulff M, Moffat K (1996) Photolysis of the carbon monoxide complex of myoglobin: nanosecond time-resolved crystallography. Science 274(5293):1726–1729
49. Lim MH, Jackson TA, Anfinrud PA (1995) Midinfrared vibrational-spectrum of co after photodissociation from heme evidence for a ligand docking site in the heme pocket of hemoglobin and myoglobin. J Chem Phys 102(11):4355–4366
50. Lim M, Jackson TA, Anfinrud PA (1995) Binding of co to myoglobin from a heme pocket docking site to form nearly linear FE-C-O. Science 269(5226):962–966
51. Franzen S, Bohn B, Poyart C, Martin JL (1995) Evidence for subpicosecond heme doming in hemoglobin and myoglobin – a time-resolved resonance Raman comparison of carbonmonoxy and deoxy species. Biochemistry 34(4):1224–1237
52. Srajer V, Ren Z, Teng TY, Schmidt M, Ursby T, Bourgeois D, Pradervand C, Schildkamp W, Wulff M, Moffat K (2001) Protein conformational relaxation and ligand migration in myoglobin: a nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction. Biochemistry 40(46):13802–13815
53. Jung YO, Lee JH, Kim J, Schmidt M, Moffat K, Šrajer V, Ihee H (2013) Volume-conserving trans–cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. Nat Chem 5(3):212–220
54. Schotte F, Cho HS, Kaila VRI, Kamikubo H, Dashdorj N, Henry ER, Graber TJ, Henning R, Wulff M, Hummer G, Kataoka M, Anfinrud PA (2012) Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. Proc Natl Acad Sci 109 (47):19256–19261
55. Ren Z, Perman B, Šrajer V, Teng TY, Pradervand C, Bourgeois D, Schotte F, Ursby T, Kort R, Wulff M, Moffat K (2001) A molecular movie at 1.8 Å resolution displays the photocycle of photoactive yellow protein, a eubacterial blue-light receptor, from nanoseconds to seconds. Biochemistry 40(46):13788–13801
56. Helliwell JR, Mitchell EP (2015) Synchrotron radiation macromolecular crystallography: science and spin-offs. Iucrj 2:283–291
57. Moffat K (2014) Time-resolved crystallography and protein design: signalling photoreceptors and optogenetics. Philos Trans R Soc B Biol Sci 369(1647):20130568
58. Moffat K (2001) Time-resolved biochemical crystallography: a mechanistic perspective. Chem Rev 101(6):1569–1581
59. Schmidt M (2019) Time-resolved macromolecular crystallography at pulsed X-ray sources. Int J Mol Sci 20(6)
60. Nam KH (2019) Sample delivery Media for Serial Crystallography. Int J Mol Sci 20(5)
61. Chapman HN (2019) X-ray free-electron lasers for the structure and dynamics of macromolecules. In: Kornberg RD (ed) Annual review of biochemistry, vol 88. Annual Reviews, pp 35–58
62. Srajer V, Schmidt M (2017) Watching proteins function with time-resolved x-ray crystallography. J Phys D Appl Phys 50(37)
63. Johansson LC, Stauch B, Ishchenko A, Cherezov V (2017) A bright future for serial femtosecond crystallography with XFELs. Trends Biochem Sci 42(9):749–762

64. Jaeger K, Dworkowski F, Nogly P, Milne C, Wang M, Standfuss J (2016) Serial millisecond crystallography of membrane proteins. In: Moraes I (ed) Next generation in membrane protein structure determination. Advances in experimental medicine and biology, vol 922, pp 137–149
65. Dickerson JL, McCubbin PTN, Garman EF (2020) RADDOSE-XFEL: femtosecond time-resolved dose estimates for macromolecular X-ray free-electron laser experiments. J Appl Cryst 53:549–560
66. de la Mora E, Coquelle N, Bury CS, Rosenthal M, Holton JM, Carmichael I, Garman EF, Burghammer M, Colletier JP, Weik M (2020) Radiation damage and dose limits in serial synchrotron crystallography at cryo- and room temperatures. Proc Natl Acad Sci U S A 117 (8):4142–4151
67. Garman EF, Weik M (2019) X-ray radiation damage to biological samples: recent progress. J Synchrotron Radiat 26:907–911
68. Lee D, Park S, Lee K, Kim J, Park G, Nam KH, Baek S, Chung WK, Lee JL, Cho Y, Park J (2020) Application of a high-throughput microcrystal delivery system to serial femtosecond crystallography. J Appl Cryst 53:477–485
69. Cohen MD, Schmidt GMJ, Sonntag FI (1964) Topochemistry. Part 2. Photochemistry of trans-cinnamic acids. J Chem Soc:2000
70. Cohen MD, Schmidt GMJ, Flavian S (1964) Topochemistry. Part 6. Experiments on photochromy + thermochromy of crystalline anils of salicylaldehydes. J Chem Soc:2041
71. Cohen MD, Schmidt GMJ (1964) Topochemistry. Part 1. Survey. J Chem Soc:1996
72. Cohen MD (1975) The photochemistry of organic solids. Angew Chem Int Ed Engl 14 (6):386–393
73. Natarajan A, Bhogala BR (2011) Bimolecular photoreactions in the crystalline state. Wiley, Hoboken, pp 175–228
74. Jarvis AG, Sparkes HA, Tallentire SE, Hatcher LE, Warren MR, Raithby PR, Allan DR, Whitwood AC, Cockett MCR, Duckett SB, Clark JL, Fairlamb IJS (2012) Photochemical-mediated solid-state 2+2 -cycloaddition reactions of an unsymmetrical dibenzylidene acetone (monothiophos-dba). CrstEngComm 14(17):5564–5571
75. Cao DK, Sreevidya TV, Botoshansky M, Golden G, Benedict JB, Kaftory M (2010) Kinetics of solid state photodimerization of 1,4-dimethyl-2-pyridinone in its molecular compound. J Phys Chem A 114(27):7377–7381
76. Benedict JB, Coppens P (2009) Kinetics of the single-crystal to single-crystal two-photon photodimerization of alpha-trans-cinnamic acid to alpha-truxillic acid. J Phys Chem A 113 (13):3116–3120
77. Zhang JJ, Zou Q, Tian H (2013) Photochromic materials: more than meets the eye. Adv Mater 25(3):378–399
78. Cole JM (2011) A new form of analytical chemistry: distinguishing the molecular structure of photo-induced states from ground-states. Analyst 136(3):448–455
79. Manrique-Juarez MD, Rat S, Salmon L, Molnar G, Quintero CM, Nicu L, Shepherd HJ, Bousseksou A (2016) Switchable molecule-based materials for micro- and nanoscale actuating applications: achievements and prospects. Coord Chem Rev 308:395–408
80. Gural'skiy IA, Quintero CM, Costa JS, Demont P, Molnar G, Salmon L, Shepherd HJ, Bousseksou A (2014) Spin crossover composite materials for electrothermomechanical actuators. J Mater Chem C 2(16):2949–2955
81. Shepherd HJ, Gural'skiy IA, Quintero CM, Tricard S, Salmon L, Molnár G, Bousseksou A (2013) Molecular actuators driven by cooperative spin-state switching. Nat Commun 4 (1):2607
82. Coppens P, Gerlits O, Vorontsov II, Kovalevsky AY, Chen YS, Graber T, Gembicky M, Novozhilova IV (2004) A very large Rh-Rh bond shortening on excitation of the [Rh-2 (1,8-diisocyano-p-menthane)(4)](2+) ion by time-resolved synchrotron X-ray diffraction. Chem Commun 19:2144–2145
83. Kim CD, Pillet S, Wu G, Fullagar WK, Coppens P (2002) Excited-state structure by time-resolved X-ray diffraction. Acta Crystallogr A 58:133–137

84. Schaniel D, Casaretto N, Bendeif E, Woike T, Gallien AKE, Klufers P, Kutniewska SE, Kaminski R, Bouchez G, Boukheddaden K, Pillet S (2019) Evidence for a photoinduced isonitrosyl isomer in ruthenium dinitrosyl compounds. CrstEngComm 21(38):5804–5810
85. Cole JM, Velazquez-Garcia JJ, Gosztola DJ, Wang SG, Chen Y-S (2018) η2-SO2 linkage photoisomer of an osmium coordination complex. Inorg Chem 57(5):2673–2677
86. Cole JM (2004) Single-crystal X-ray diffraction studies of photo-induced molecular species. Chem Soc Rev 33(8):501–513
87. Coppens P, Novozhilova I, Kovalevsky A (2002) Photoinduced linkage isomers of transition-metal nitrosyl compounds and related complexes. Chem Rev 102(4):861–883
88. Cormary B, Ladeira S, Jacob K, Lacroix PG, Woike T, Schaniel D, Malfant I (2012) Structural influence on the photochromic response of a series of ruthenium mononitrosyl complexes. Inorg Chem 51(14):7492–7501
89. Zangl A, Klufers P, Schaniel D, Woike T (2009) Photoinduced linkage isomerism of binuclear bis(pyrazole-3,5-dicarboxylato)-bridged {RuNO}(6) centres. Inorg Chem Commun 12 (10):1064–1066
90. Dieckmann V, Eicke S, Rack JJ, Woike T, Imlau M (2009) Pronounced photosensitivity of molecular Ru(bpy)(2)(OSO) (+) solutions based on two photoinduced linkage isomers. Opt Express 17(17):15052–15060
91. Schaniel D, Bendeif EE, Woike T, Bottcher HC, Pillet S (2018) Wavelength-selective photoisomerisation of nitric oxide and nitrite in a rhodium complex. CrstEngComm 20 (44):7100–7108
92. Galle G, Nicoul M, Woike T, Schaniel D, Freysz E (2012) Unraveling the mechanism of NO ligand photoisomerism by time-resolved infrared spectroscopy. Chem Phys Lett 552:64–68
93. Schaniel D, Woike T (2009) Necessary conditions for the photogeneration of nitrosyl linkage isomers. Phys Chem Chem Phys 11(21):4391–4395
94. Coppens P, Zheng SL, Gembicky M (2008) Static and time-resolved photocrystallographic studies in supramolecular solids. Z Kristall 223(4–5):265–271
95. Phillips AE, Cole JM, d'Almeida T, Low KS (2012) Ru–OSO coordination Photogenerated at 100 K in Tetraammineaqua(sulfur dioxide)ruthenium(II) (±)-Camphorsulfonate. Inorg Chem 51(3):1204–1206
96. Hatcher LE (2016) Raising the (metastable) bar: 100% photo-switching in [Pd(Bu4dien)(η1-O2)]+ approaches ambient temperature. CrstEngComm 18(22):4180–4187
97. Ohashi Y, Yanagi K, Kurihara T, Sasada Y, Ohgo Y (1981) Crystalline state reaction of cobaloxime complexes by x-ray exposure. 1. Direct observation of cobalt-carbon bond cleavage in [(R)-1-cyanoethyl][(S)-(−)-.alpha.-methylbenzylamine]bis(dimethylglyoximato)-cobalt(III). J Am Chem Soc 103(19):5805–5812
98. Decurtins S, Gutlich P, Kohler CP, Spiering H, Hauser A (1984) Light-induced excited spin STATE trapping in a transition-metal complex – the hexa-1-propyltetrazole-iron (II) tetrafluoroborate-crossover system. Chem Phys Lett 105(1):1–4
99. Carbonera C, Costa JS, Money VA, Elhaik J, Howard JAK, Halcrow MA, Letard JF (2006) Photomagnetic properties of iron(II) spin crossover complexes of 2,6-dipyrazolylpyridine and 2,6-dipyrazolylpyrazine ligands. Dalton Trans 25:3058–3066
100. Thompson AL, Money VA, Goeta AE, Howard JAK (2005) Structural studies of thermal- and light-induced transitions in iron(II) spin-crossover complexes. C R Chim 8(9–10):1365–1373
101. Money VA, Evans IR, Elhaik J, Halcrow MA, Howard JAK (2004) An X-ray powder diffraction study of the spin-crossover transition and structure of bis (2,6-dipyrazol-1-ylpyrazine)iron(II) perchlorate. Acta Crystallogr B Struct Sci Cryst Eng Mater 60:41–45
102. Money VA, Elhaik J, Halcrow MA, Howard JAK (2004) The thermal and light induced spin transition in FeL2 (BF4)(2) (L=2,6-dipyrazol-1-yl-4-hydroxymethylpyridine). Dalton Trans 10:1516–1518
103. Money VA, Elhaik J, Evans IR, Halcrow MA, Howard JAK (2004) A study of the thermal and light induced spin transition in FeL2 (BF4)(2) and FeL2 (ClO4)(2) L=2,6-di(3-methylpyrazol-1-yl)pyrazine. Dalton Trans 1:65–69

104. Money VA, Costa JS, Marcen S, Chastanet G, Elhaik J, Halcrow MA, Howard JAK, Letard JF (2004) A photomagnetic study of three iron(II) compounds containing ligands from the 2,6-di (pyrazol-1-yl)pyridine series. Chem Phys Lett 391(4–6):273–277
105. Letard JF, Guionneau P, Rabardel L, Howard JAK, Goeta AE, Chasseau D, Kahn O (1998) Structural, magnetic, and photomagnetic studies of a mononuclear iron(II) derivative exhibiting an exceptionally abrupt spin transition. Light-induced thermal hysteresis phenomenon. Inorg Chem 37(17):4432–4441
106. Gutlich P, Gaspar AB, Garcia Y (2013) Spin state switching in iron coordination compounds. Beilstein J Org Chem 9:342–391
107. Halcrow MA (2008) Trapping and manipulating excited spin states of transition metal compounds. Chem Soc Rev 37(2):278–289
108. Techert S, Schotte F, Wulff M (2001) Picosecond x-ray diffraction probed transient structural changes in organic solids. Phys Rev Lett 86(10):2030–2033
109. Rice SF, Gray HB (1983) Electronic absorption and emission spectra of binuclear platinum (II) complexes. Characterization of the lowest singlet and triplet excited states of tetrakis (diphosphonato)diplatinate(4-) anion (Pt2(H2P2O5)44-). J Am Chem Soc 105(14):4571–4575
110. Leung KH, Phillips DL, Che CM, Miskowski VM (1999) Resonance Raman intensity analysis investigation of metal-metal bonded transitions: an examination of the (1)A(2u) <– (1)A (1g) (5d sigma* -> 6p sigma) transition of Pt-2(P2O5H2)(4)(4-). J Raman Spectrosc 30 (11):987–993
111. Thiel DJ, Liviņs P, Stern EA, Lewis A (1993) Microsecond-resolved XAFS of the triplet excited state of Pt2(P2O5H2)4(4-). Nature 362(6415):40–43
112. Ozawa Y, Terashima M, Mitsumi M, Toriumi K, Yasuda N, Uekusa H, Ohashi Y (2003) Photoexcited crystallography of diplatinum complex by multiple-exposure IP method. Chem Lett 32(1):62–63
113. Christensen M, Haldrup K, Bechgaard K, Feidenhans'l R, Kong QY, Cammarata M, Lo Russo M, Wulff M, Harrit N, Nielsen MM (2009) Time-resolved X-ray scattering of an electronically excited state in solution. Structure of the (3)A(2u) state of Tetrakis-mu-pyrophosphitodiplatinate(II). J Am Chem Soc 131(2):502–508
114. van der Veen RM, Milne CJ, Pham VT, El Nahhas A, Weinstein JA, Best J, Borca CN, Bressler C, Chergui M (2008) EXAFS structural determination of the Pt-2(P2O5H2)(4)(4-) anion in solution. Chimia 62(4):287–290
115. Miskowski VM, Rice SF, Gray HB, Dallinger RF, Milder SJ, Hill MG, Exstrom CL, Mann KR (1994) Spectroscopy and photophysics of RH2(dimen)4(2+) (dimen = 1,8-diisocyanomenthane) – exceptional metal-metal bond shortening in the lowest electronic excited-states. Inorg Chem 33(13):2799–2807
116. Vorontsov II, Kovalevsky AY, Chen YS, Graber T, Gembicky M, Novozhilova IV, Omary MA, Coppens P (2005) Shedding light on the structure of a photoinduced transient excimer by time-resolved diffraction. Phys Rev Lett 94(19):4
117. Vorontsov II, Graber T, Kovalevsky AY, Novozhilova IV, Gembicky M, Chen YS, Coppens P (2009) Capturing and analyzing the excited-state structure of a Cu(I) phenanthroline complex by time-resolved diffraction and theoretical calculations. J Am Chem Soc 131(18):6566–6573
118. Ohashi Y (2013) Dynamic motion and various reaction paths of cobaloxime complexes in crystalline-state photoreaction. Crystallogr Rev 19(Suppl 1):2–146
119. Coppens P, Zheng SL, Gembicky M, Messerschmidt M, Dominiak PM (2006) Supramolecular solids and time-resolved diffraction. CrstEngComm 8(10):735–741
120. Zheng SL, Gembicky M, Messerschmidt M, Dominiak PM, Coppens P (2006) Effect of the environment on molecular properties: synthesis, structure, and photoluminescence of Cu(I) bis (2,9-dimethyl-1,10-phenanthroline) nanoclusters in eight different supramolecular frameworks. Inorg Chem 45(23):9281–9289
121. Zheng SL, Messerschmidt M, Coppens P (2005) An unstable ligand-unsupported Cu-I dimer stabilized in a supramolecular framework. Angew Chem Int Ed 44(29):4614–4617

122. Inokuma Y, Kawano M, Fujita M (2011) Crystalline molecular flasks. Nat Chem 3 (5):349–358
123. Kawano M, Fujita M (2007) Direct observation of crystalline-state guest exchange in coordination networks. Coord Chem Rev 251(21–24):2592–2605
124. Kawano M, Kobayashi Y, Ozeki T, Fujita M (2006) Direct crystallographic observation of a coordinatively unsaturated transition-metal complex in situ generated within a self-assembled cage. J Am Chem Soc 128(20):6558–6559
125. Haneda T, Kawano M, Kawamichi T, Fujita M (2008) Direct observation of the labile imine formation through single-crystal-to-single-crystal reactions in the pores of a porous coordination network. J Am Chem Soc 130(5):1578–1579
126. Kawamichi T, Haneda T, Kawano M, Fujita M (2009) X-ray observation of a transient hemiaminal trapped in a porous network. Nature 461(7264):633–635
127. Horiuchi S, Murase T, Fujita M (2011) Noncovalent trapping and stabilization of dinuclear ruthenium complexes within a coordination cage. J Am Chem Soc 133(32):12445–12447
128. Takezawa H, Murase T, Fujita M (2012) Temporary and permanent trapping of the metastable twisted conformer of an overcrowded chromic alkene via encapsulation. J Am Chem Soc 134 (42):17420–17423
129. Blake AJ, Champness NR, Easun TL, Allan DR, Nowell H, George MW, Jia J, Sun XZ (2010) Photoreactivity examined through incorporation in metal-organic frameworks. Nat Chem 2 (8):688–694
130. Easun TL, Jia JH, Reade TJ, Sun XZ, Davies ES, Blake AJ, George MW, Champness NR (2014) Modification of coordination networks through a photoinduced charge transfer process. Chem Sci 5(2):539–544
131. Makal A, Benedict J, Trzop E, Sokolow J, Fournier B, Chen Y, Kalinowski JA, Graber T, Henning R, Coppens P (2012) Restricted photochemistry in the molecular solid state: structural changes on photoexcitation of Cu(I) phenanthroline metal-to-ligand charge transfer (MLCT) complexes by time-resolved diffraction. J Phys Chem A 116(13):3359–3365
132. Kalinowski JA, Fournier B, Makal A, Coppens P (2012) The LaueUtil toolkit for Laue photocrystallography. II. Spot finding and integration. J Synchrotron Radiat 19:637–646
133. Kaminski R, Graber T, Benedict JB, Henning R, Chen YS, Scheins S, Messerschmidt M, Coppens P (2010) Optimizing the accuracy and precision of the single-pulse Laue technique for synchrotron photo-crystallography. J Synchrotron Radiat 17:479–485
134. Jarzembska KN, Kaminski R, Fournier B, Trzop E, Sokolow JD, Henning R, Chen Y, Coppens P (2014) Shedding light on the photochemistry of coinage-metal phosphorescent materials: a time-resolved Laue diffraction study of an Ag-I-Cu-I tetranuclear complex. Inorg Chem 53(19):10594–10601
135. Jarzembska KN, Hapka M, Kaminski R, Bury W, Kutniewska SE, Szarejko D, Szczesniak MM (2019) On the nature of luminescence thermochromism of multinuclear copper (I) benzoate complexes in the crystalline state. Crystals 9(1)
136. Collet E, Lorenc M, Cammarata M, Guérin L, Servol M, Tissot A, Boillot M-L, Cailleau H, Buron-Le Cointe M (2012) 100 picosecond diffraction catches structural transients of laser-pulse triggered switching in a spin-crossover crystal. Chemistry 18(7):2051–2055
137. Bertoni R, Cammarata M, Lorenc M, Matar SF, Létard J-F, Lemke HT, Collet E (2015) Ultrafast light-induced spin-state trapping photophysics investigated in Fe(phen)2(NCS)2 spin-crossover crystal. Acc Chem Res 48(3):774–781
138. Bertoni R, Lorenc M, Laisney J, Tissot A, Moreac A, Matar SF, Boillot ML, Collet E (2015) Femtosecond spin-state photo-switching dynamics in an Fe-III spin crossover solid accompanied by coherent structural vibrations. J Mater Chem C 3(30):7792–7801
139. Coppens P (2015) Perspective: on the relevance of slower-than-femtosecond time scales in chemical structural-dynamics studies. Struct Dyn 2(2)
140. Basuroy K, Chen Y, Sarkar S, Benedict J, Coppens P (2017) Exploring the structural changes on excitation of a luminescent organic bromine-substituted complex by in-house time-resolved pump-probe diffraction. Struct Dyn 4(2)

Index