

Changing and challenging times for service crystallography†

Simon J. Coles* and Philip A. Gale*

Received 21st November 2011, Accepted 8th December 2011

DOI: 10.1039/c2sc00955b

Crystallography is no longer solely the preserve of the specialist, a situation that has implications for the operation of crystallographic service facilities. This mini-review provides an overview of the challenges in operating a crystallographic facility from the perspective and experience of the UK National Crystallography Service – a modern mid-range facility. Examples of chemical research generating the greatest challenges for the modern crystallography service and the state-of-the-art tools, hardware, facilities and expertise that are required to address them are highlighted. An overview of current research trends in single crystal diffraction research, which will ensure the future development of the technique, is presented. The remit of the service crystallographer is examined, together with proposed examples of best practice.

Introduction

Work on projects in structural chemistry in the 21st century can pose significant challenges for the chemical crystallographer.¹ It

Chemistry, University of Southampton, Southampton, UK. E-mail: s.j.coles@soton.ac.uk; philip.gale@soton.ac.uk; Fax: +44 (0)2380596723; Tel: +44 (0)2380596721

† Electronic supplementary information (ESI) available: Datasets and experimental protocols are provided for data collected on the Diamond I19 beamline, the VHF and UHF rotating anode diffractometers and a sealed tube source (image plate) diffractometer are provided. These a) serve as exemplars for data collected at the NCS and b) form the basis for a comparison of the power and capability of the different instruments. CCDC reference numbers 855293–855295. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c2sc00955b

is no longer the case that a structure determination is performed only if the sample lends itself well to the experiment. A project may rely on a structure determination, and it may be imperative to get the most information out of a crystal, no matter what its size or quality. Moreover, the purpose of a crystal structure determination has changed over recent years. Not only may there be a requirement for characterisation of a molecular structure, but also to investigate whole crystal structures – that is the ensemble of molecules in the lattice – to address questions, for example, in the areas of crystal engineering and supramolecular chemistry,² polymorphism³ and structural systematics.⁴ In addition, there is an increasing realisation that crystal structures are not static; photocrystallography⁵ and high-pressure⁶ studies are well established techniques. Finally, we are in an ‘informatics era’ where, due to the volume of data amassed, it is possible to



Simon J Coles

Simon J. Coles is Operations Director of the National Crystallography Service and a Lecturer at the University of Southampton. He started working with the National Crystallography Service in 1992 and became its manager in 1998 and Director in 2009. His research interests include structure–property systematics, phase transition characterisation, charge density determination, crystal growth and the application of eResearch techniques to scientific endeavour.



Philip A Gale

Philip A. Gale is Professor of Supramolecular Chemistry and Head of Chemistry at the University of Southampton. In May 2010 he assumed the role of Head of the UK’s National Crystallography Service acting as the ‘User at the Centre’ ensuring that the service continues to evolve to meet the changing needs of the UK research community. His research interests focus on the supramolecular chemistry of anionic species in particular the recognition, sensing and lipid bilayer transport of anions.

perform new science by asking meaningful questions of structural databases in order to exert control over self-assembly processes in the solid state⁷ and to predict crystal structures of new molecular systems.⁸

Work involving samples that are too small or diffract too weakly for data collection on a conventional laboratory diffractometer will require recourse to the most highly powered instrumentation available. Cost implications may then require the establishment of shared facilities. This has been achieved by the establishment, in the UK and some other countries, of national services and access to national or international synchrotron facilities. The UK National Crystallography Service (NCS) has been in continuous operation since its founding in 1980 by Hursthouse.⁹ During this time, the NCS has pioneered several aspects of hardware development through close collaboration with instrument manufacturers. Moreover, the university laboratories in which the NCS has operated have always provided departmental support and conducted chemical crystallography research. Effective management of these modes of operation has enabled the facility to generate a particularly high throughput of samples. Drawing on this experience, the new generation of the NCS (funded from May 2010) has identified that, to support chemical crystallography on a national scale, there is a requirement for a) the use of powerful state-of-the-art instrumentation and b) a specific approach towards handling the large volumes of data arising from a high turnover of samples. Furthermore, an increased understanding on the part of crystal providers as to the kind of results crystallography can offer, given the limitations of *e.g.* crystal quality and content, will increase the success rate of experiments.

This review reflects on these objectives and experiences and provides some insight into developments that can move the field of service crystallography forward. To this end, we highlight the current challenges facing a mid-range service, showing how these can be met and demonstrating how these efforts also contribute to the future development of the technique in a more conventional laboratory.

Background – instrumentation and staffing the national crystallography service

The UK NCS in its present incarnation was the first service funded under the new UK Engineering and Physical Science Research Council (EPSRC) mid-range facility programme,¹⁰ which provides shared infrastructure supporting projects across the entire Research Council remit. The service is available at no cost to those eligible to apply to the EPSRC for research funding (subject to approval by the service's Management Access Panel). Given the position such a facility holds in the community, it can act as an exemplar for other mid-range facilities and any similar service operating at a medium, large or national scale. Such a facility may often uncover challenges and issues that will be met at a later stage by smaller scale operations and there is merit in considering these for the benefit of the wider community in the future.

The NCS provides a two-centred, three-tiered model where synthetic chemistry researchers and crystallographers have recourse to centralised high-powered laboratory and subsequent synchrotron facilities that can tackle more challenging problems

than are possible with more typical homelab instrumentation. In the three-tiered model, departmental services filter out the samples they can examine and pass on those that are too small or weakly scattering to the NCS. This centralized, laboratory-based high intensity facility then further screens and filters out another level of samples and acts as a funnel for the synchrotron (Diamond Light Source). Thus samples are only ever examined on the most suitable instrumentation and only those really requiring expensive and oversubscribed synchrotron facilities are studied there. The NCS also provides support for researchers with no local facilities in the technique and can be for use in case of emergency for other crystallography groups (*e.g.* local equipment breakdown).

The recently installed equipment base at the NCS comprises the most powerful molybdenum-based rotating anode source currently available, which has been customized from the Cu-based Rigaku FR-E+ SuperBright. This source is equipped with two kappa geometry diffractometers, using either a VariMax-VHF or a VariMax HF optics developed for Mo radiation and designed to provide complementary beam profiles. Each of these instruments are furnished with a state-of-the-art Saturn 724 + HG CCD detector. The FR-E+ equipment forms the primary instrument base of the NCS, but this is also supplemented and complemented by access to both a high intensity Cu-based CCD system funded by an institutional initiative, aimed at supporting the integration of chemical and biological structural crystallography, and a Mo-based sealed tube – image plate (SPIDER). The Diamond beamline I19¹¹ is the synchrotron facility of choice for the small molecule service crystallographer, being one of a small number of dedicated and optimised beamlines in the world for this technique. The beamline consists of two experimental hutches, with EH1 being designed for service work and EH2 being dedicated to subjecting samples to extreme conditions. With highly advanced and tailored beam conditioning specifications, EH1 collects data on the most challenging samples with a bespoke kappa geometry goniometer and a Saturn 724 + CCD detector. Of particular note is the availability of robotic automation of the mounting process¹² for screening these challenging samples, where numerous crystals can be tried, ranked and revisited in order to assess the most appropriate for data collection. A full description of experimental approach, details of data collection protocols and exemplar datasets from the synchrotron, rotating anode and sealed tube instruments outlined above are provided in the Electronic Supplementary Information†.

An operation of this size requires significant staffing - four post-doctoral researchers and further technical and administrative support - to optimise sample throughput. It is also important for an operation of this nature to devote a certain amount of time to performing research that can feed into the development of the service, maintaining its cutting edge and moving the subject forward. For example, in the case of the NCS this type of research is based around projects involving a) systematic study of the crystal structures of large families of related compounds, addressing issues such as polymorphism and structural similarity, and b) charge density studies, answering questions surrounding the nature of chemical bonding and its influence on crystal structure and reactivity. Whilst providing valuable chemical insight, this research program also moves crystallography forward by developing thinking and methodology towards

data management, comparison of large numbers of *complete* crystal structures and pushing the limits of collecting high resolution data towards a goal of routinely examining structures at an electronic, rather than atomic, level.

Current challenges in service crystallography

With every leap in technology a new set of challenges are quickly found. There have been several step-changes for service crystallography, notably the introduction of automated diffractometers in the 1960's and 70's and the dramatic increase in computing accessibility and power in the last couple of decades. It was some 20 years ago that area detectors were adopted by the chemical crystallography community¹³ and since then some significant, but relatively incremental, advances have been made. Arguably, we are nearing the end of this era and are addressing new challenges with relatively old technology.

Many areas of modern research are producing increasing numbers of poor quality crystals, due to inherent characteristics relating to their chemistry and this has implications for the collection, processing and work-up of data and the quality of the final result. Accordingly, certain classes of compounds of current interest *e.g.* metal-organic frameworks or biologically relevant supramolecular complexes, tend to form predominantly small or weakly scattering crystals. Modern synthetic procedures are capable of producing large and sophisticated systems – much like working with protein structures – and these systems generally have a high degree of conformational flexibility, leading to disorder and lower resolution datasets. Many crystals contain a significant percentage of solvent or void space. These solvent molecules tend to be poorly ordered and produce diffuse diffraction patterns of limited resolution.

Crystallisation is a subtle, complex and poorly understood process, but it is clear that nucleation and crystal growth are influenced by a number of factors.¹⁴ It is not simply a case of controlling evaporation rate or understanding phase diagrams relating factors such as temperature and concentration. Other factors, such as the nature of nucleation sites or presence of impurities, can affect crystal quality or form. The quality of a crystalline sample affects the quality of the resultant crystal structure, yet more often than not little consideration is given to the crystallisation step irrespective of the amount of time and money invested in the initial creation of the product. The consequence is that crystallography services are increasingly being challenged by large numbers of samples and an expectation of a quick turnaround time, as structure determination can often be as fast as alternative characterisation techniques.

With the introduction and development of automated and then computer controlled instrumentation, crystallography became an increasingly accessible technique from the 1960's onwards and many academics were active in this field. It is this generation of researchers that drove the subject forward, but are now retiring and not being replaced as rapidly with new appointments. Accordingly, whilst software has become more usable, there have been relatively few innovations or new algorithms developed in the last decade to deal with new challenges.

There has been an increase in computing power recently such that, without complete end-to-end automation, it is difficult to reduce the time taken to solve and refine crystal structures.

Moreover, at the current rate of generation it is no longer possible to keep up with the dissemination of results under the constraints of the conventional publishing processes. This situation, coupled with the fact that publication of crystallographic data is generally tied to the publication of a journal article describing its synthesis, means that only a fraction of the crystal structures determined are reaching the public domain.

Data management and preservation of digital information over time and across different storage media is becoming challenging in a fast-developing world – those labs that have been active for the last 20 years will probably have archives based on at least four different media, and managing and storing these records is becoming an increasingly involved task. As the throughput of crystallographic results generation increases, this situation will be greatly exacerbated and will hamper the advance and progress of structural chemistry. Additionally, the funding agencies are beginning to stipulate that the outputs from the work they support, which are often taken to be data in addition to conventional journal articles, must be appropriately managed and accessible.¹⁵ These policies on access to research outputs will have a profound impact on the way in which research is performed and how results are communicated.

Addressing these challenges

It is clear that structure determination by single crystal diffraction is a valuable technique that underpins chemical science and is often a research theme in its own right. Some of the current pressing issues of concern to modern chemical crystallography may be broadly grouped into the following categories:

- i) Instrumentation development, control and interaction;
- ii) Training and education;
- iii) Size and coordination of the software development community;
- iv) Managing and using the increase in data.

Instrumentation

In the last decade, the nature and complexity of the products of synthesis in chemistry has increased significantly, whilst service crystallographers supporting this work are generally doing so with a mature technology developed in the 1990's. This can be addressed by using state-of-the-art instrumentation and proactively partaking in its continuing development and also that of its control software. Through a close relationship with an instrument manufacturer, the brightest molybdenum-source X-ray generator has been designed and constructed, giving an intensity in the home laboratory that is of a similar order to that of a second generation synchrotron. The exceptionally high-powered rotating anode X-ray generator is coupled to a high sensitivity CCD detector – a combination enabling fast data collection times and the ability to examine extremely small and weakly diffracting crystals. Recourse to large-scale centralised facilities and the appropriate staffing to manage the beamtime and arising data is invaluable and an appropriate way to address the requirements of modern chemistry. A 'two-centred, three-tiered' model (1 = "home laboratory"; 2 = "National Service laboratory"; 3 = "synchrotron facility"), whereby weak diffractors have been screened on the most powerful laboratory source available

before being referred to the synchrotron, ensures that the right samples are matched to the correct facility. A periodic allocation of beamline time ensures that only appropriate samples are examined in a timely manner on this oversubscribed and expensive-to-run facility. We believe that the diffraction limit is on the verge of being reached for particular types of chemical systems, as severe radiation damage is observed for approximately 8% of samples investigated at a third generation synchrotron. It will soon be necessary to adopt the approaches that the protein crystallography community have taken to address this problem.¹⁶

Experiences of operating over both the laboratory and synchrotron facility enables a comparison between the two to be made, thereby providing an understanding of what can be achieved on such modern equipment. Table 1 outlines a comparison of a data collection on the same crystal under experimental conditions designed to be almost identical. In this experiment, data were collected at a synchrotron (1 s individual image exposure time, 70% attenuated) and on a rotating anode (10 and 30 s individual image exposures).

In summary, the quality of the results from these comparison data collections demonstrate that a routine data collection with images collected at 1 s each on the attenuated synchrotron is approximately equivalent to a 30 s per image data collection on a state-of-the-art rotating anode. For example, the final *R* factors are 3.45% and 3.25% for data merging with values of 8.1% and 8.55% for the structure refinements. When taking into account the difference in time factor and attenuation applied, the synchrotron is collecting comparable data 100 times faster. However, a data collection with a 10 s per image exposure time

on the rotating anode still produces a publishable dataset, albeit of slightly lower quality (the synchrotron comparison here is 30 times faster when attenuation is considered). Interestingly, the same sample could only be investigated on a (graphite monochromated) sealed tube equipped with an image plate detector due to the unusually long exposures required – a comparable unit cell could only be determined from 12 h exposures (which equates to a 5 month data collection!). It is interesting to compare these results with those of Stalke *et al.*,¹⁷ who made a comparison between sealed tube microfocus and rotating anode laboratory-based sources.

A thermal ellipsoid plot of the resulting structure is shown below (Fig. 1, ellipsoids drawn at 30% probability level), as obtained from the 10 s dataset collected on the rotating anode diffractometer. There is some minor disorder of the tetramethylammonium cation, but this was not included in the model. CIF files for each of the three compared full data collections are provided as Electronic Supporting Information†.

CCD detectors have been in common use for over 15 years and are arguably reaching the technical limit of their capabilities. New approaches to the measurement of diffracted intensities, such as Hybrid Pixel Detectors based on CMOS technology¹⁸ will allow for faster and more sensitive data collection. The widespread use of focusing mirrors for chemical crystallography¹⁹ has revolutionised X-ray generator technology for the routine service laboratory, particularly through microfocus sources.²⁰ New approaches through the use of liquid gallium as a target²¹ are particularly promising for improving the dissipation of heat and thereby producing higher powered X-rays. These new technologies will also be less demanding and more

Table 1 Comparative data collection and refinement parameters for the tetramethylammonium acetate complex of a bis-3,5-dinitrophenyl urea compound collected on beamline I19 of the Diamond Light Source and the rotating anode source at the NCS (a full dataset was not collected on the sealed tube instrument due to the length of time it would have taken)

Instrument	Synchrotron	Rotating Anode	Rotating Anode
Exposure	1 s	10 s	30 s
$\lambda/\text{Å}$	0.6889	0.71073	0.71073
<i>T/K</i>		293(2)	
Sum formula		$\text{C}_{19}\text{H}_{23}\text{N}_7\text{O}_{11}$	
Formula weight		525.44	
Crystal size/mm		0.07 \times 0.05 \times 0.02	
Crystal system		Triclinic	
Space group		$\text{P}\bar{1}$	
<i>Z</i>		2	
<i>a</i> /Å	9.537(7)	9.510(17)	9.525(12)
<i>b</i> /Å	11.248(9)	11.31(2)	11.380(15)
<i>c</i> /Å	12.439(10)	12.46(2)	12.488(16)
α (°)	107.701(9)	107.583(19)	107.619(15)
β (°)	97.463(6)	97.533(14)	97.507(10)
γ (°)	104.937(10)	104.719(19)	104.894(16)
<i>V</i> /Å ³	1196.4(16)	1203(4)	1214(3)
Reflections used in unit cell refinement	2318	2974	2547
θ range for unit cell	1.9–26.6	3.0–27.5	3.0–27.4
All data	11746	11790	11747
θ range for all data [°]	2.43–26.60	2.96–27.48	2.95–27.46
Unique data	5223	5276	5265
<i>R</i> _{int}	0.0345	0.0389	0.0325
Observed data	3330	3619	3188
<i>R</i> ₁ , <i>wR</i> ₂ [<i>I</i> > 2 <i>σ</i> (<i>I</i>)]	0.0810/0.2357	0.1249/0.2658	0.0855/0.2300
<i>R</i> ₁ , <i>wR</i> ₂ [all data]	0.1108/0.2662	0.1709/0.2959	0.1329/0.2674
GoF	1.072	1.237	1.109
Residuals/e Å ⁻³	0.389/–0.318	0.407/–0.212	0.434/–0.254

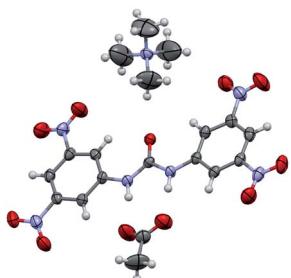


Fig. 1 A thermal ellipsoid plot (30% probability level) as obtained from the 10 s data set collected on the rotating anode diffractometer.

economical to operate than current hardware. It is key that a close relationship exists with the equipment manufacturers, allowing for enhancement of the capability of instrumentation and adaptation of hardware to tackle new science and experiments, for example investigating structure under dynamic conditions or in excited states.

Training and education

One way to keep apace of the rate at which data are being generated is to increase the number of operators in a facility. This can practically be achieved by training the providers of crystals to perform their own experiments. Software and experimentation have progressed recently to the point that non ‘classically trained’ researchers, *e.g.* PhD in chemistry, can readily perform routine crystal structure determinations. Therefore, there is a need for training programs ranging from introductory courses *e.g.* lecture module components that may be used by other institutions, to advanced skills workshops to bring experienced researchers together.²² The biennial Intensive School in crystal structure determination held at Durham University is a clear exemplar of this approach in the United Kingdom,²³ whilst the ACA school performs a similar role in the United States,²⁴ however the focus of some schools is often on the analysis of data as opposed to the practical aspects of its collection. Centralised and mid-range facilities are in an excellent position to provide hands-on experience and involve their users in the data acquisition process and, in particular, the handling and manipulation of challenging samples and the subsequent raw data integration and correction. Perhaps the most crucial aspect of training is the much overlooked area of (re)crystallisation – the cause of the vast majority of problems with data integration, structure solution and refinement can be attributed to poor crystal quality. Whilst some types of system may never produce good quality crystals, most attempts at crystal growth are made by synthetic chemists with little formal training in the understanding and importance of good quality crystals – most synthesis procedures take days or even months and crystallisation is at best a process undertaken over a much shorter period of time and usually only by solvent evaporation. There are few routes available to educate researchers in this respect and even rudimentary online resources could make a considerable difference.

An important question to ask is ‘do we need to reconsider the purpose and value of a crystal structure?’. Often a crystal structure is not considered to be of sufficient quality to merit publication in a journal article – if the purpose of the inclusion of

the structure in an article is merely as a proof of the formation of a particular product then surely establishing connectivity is all that is required and the quality of the structure can be compromised to an extent. A primary reason for upholding such standards is so that structural databases have appropriate quality data. However, if the crystallographic community can derive protocols for describing the quality of a structure then, in principle, everything can be made available (*vide infra*). Thereby a ‘data fit for purpose’ approach could be taken by those who wish to use crystal structures for further work *e.g.* as the starting point for computational studies.

Finally, one role that the NCS considers important is the engagement of the general public with the work conducted by the service. This is achieved through an outreach programme whereby the service contributes to larger events in the university and also runs dedicated events focusing on crystallography.

Software development

There is now a lack of crystallographically trained software developers. We have moved on from an era where researchers themselves would write their own software as required to solve a particular problem to one where the primary interest is in the result itself, rather than how it was derived. Consequently there is little merit or reward in developing scientific software under current recognition or assessment systems in academia. However, we are operating in an environment where computation is prevalent and pervasive, and drivers and incentives for skills to be developed in these areas must be put in place. The journal Open Research Computation²⁵ provides precisely this, whereby peer reviewed software code itself can be published and therefore developers can get the recognition they require for assessment exercises and career progression. As an alternative, or first step to preserving and making software available, CCP4²⁶ acts as an important repository for new or historic code. Additionally, there are initiatives aimed at educating the next generation of crystallographic software developers – particularly the IUCr Crystallographic Computing School organised by the IUCr Commission on Crystallographic Computing.²⁷ This is not to say that there are no new developments – recently a new approach to structure solution, Charge Flipping,²⁸ has been revolutionary and is proving to be adept at dealing with poor data. Moreover, the generation of researchers that are responsible for the majority of crystallographic software used today are approaching retirement. There is therefore a need for crystallographic software that contains algorithms derived from previous generations, but that can continue to be developed – the Olex2²⁹ and CRYSTALS³⁰ software packages are examples of crystallographic software suites that are still actively developed. It should not be overlooked that the vast majority of crystal structures produced in the last three decades have been from the ubiquitous SHELX³¹ suite of software. Initiatives such as the Collaborative Computing Projects³² are useful vehicles for developing new software and training communities in its use.

We should also consider the question ‘What are crystal structures being used for these days?’. The body of structural information we have amassed over the last four decades is now driving the new science of systematics and informatics and therefore some of the drivers for generating crystal structures

have changed in the last 5–10 years. This is in contrast to the purpose of a crystal structure determination being primarily for molecular structural characterisation, which was the main purpose of the experiment up to the 1990's. There is therefore an increasing need for software for analysis of families of crystal structures³³ and also for follow-on studies, such as energy and property calculations.

Data deluge

Over the last decade there has been significant development and interest in the role that eResearch can play in the modern crystallographic facility – particularly with respect to data management and data publication.³⁴ Whilst crystallographic data is well structured and understood and only of medium size in terms of 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. As a high-throughput facility, the diffraction laboratory at Southampton generates approximately 2500 data collections per annum – this rate provides an indication of the issues that all laboratories will need to address in the future. As part of the JISC-funded “Keeping Research Data Safe” report,³⁵ a crystallographic service was used as a case study for the analysis of issues, costs and benefits surrounding the long term preservation and curation of digital research-related data. This study was able to assign costs to the storage of crystallographic information and, in particular, the funding requirements involved in migrating data across archival systems e.g. from CDs to online or removable media. A further assessment was made of the cost of lost data (either arising from the migration process, corruption or from lack of management i.e. loss) – if one considers the effort involved in trying to perform the experiment again (including resynthesising the crystal) then the financial loss for an average facility can run into tens of thousands of pounds. As previously mentioned, the funding agencies are beginning to demand that the outputs of the work they fund should be suitably managed so that they can be exploited (by anyone) at a later date, thus giving an even further return on investment.¹⁵

As a facility open to use by a wide range of researchers across a large area, the ability to track, process and deliver a high volume of datasets is paramount – this is a capability that is important for all service crystallography laboratories. Larger scale facilities are developing information management systems to address this.³⁶ These systems can support application, access, sample submission, sample tracking and reporting, data accessibility and data management at the facility level. Finally, good data management is key to the moral (as tax payers) and increasing funder requirement to make the outputs of funded research easily available for the purpose of sharing and reuse. This is now policy for many funding agencies³⁷ and many others are following this example. At best, it is generally acknowledged that most crystallographic facilities only manage to make around 15–20% of their results available to a wider audience. The ultimate aim of any facility should be to make all of its research outputs more widely available – to this end crystal structure data repositories are now becoming available.³⁸ eCrystals³⁹ has been an integral part of NCS for making available ‘full structure only’ data since May 2010 and provides a global online approach to

making available all of the data generated during the course of a crystal structure determination. This approach is also in the process of being incorporated into the NCS information management system. This human and machine readable resource has been developed in collaboration with publishers, digital librarians, data managers and informaticians and is fully integrated into the traditional publishing and database systems in common use in areas of chemistry that rely on crystallographic characterisation. A complementary solution to this problem is provided by ReciprocalNet,⁴⁰ whereby a consortia of laboratories contribute to a distributed database of molecular structures, some of which may be made openly available.

Conclusions

Structural science is in a very strong position in the chemistry research community, but needs the development and support of the right tools and approaches to sustain this going forward. The recourse to powerful instrumentation is available and exciting hardware developments are likely to occur in the near future. However, it is also important to train new generations of software developers and instrument operators to support a growing user base. Finally, the subject is in an excellent position to lead the field of digital information management and is already demonstrating the value of reusing crystal structure data.

Acknowledgements

We thank the EPSRC for funding the UK National Crystallography Service and Diamond Light Source for access to beamline I19. In particular, we would like to extend a special thank you to Emeritus Professor Mike Hursthouse, the previous director of the NCS, for useful discussions, all members of the NCS staff past and present for their contributions over the last three decades. Additionally we would like to thank Professor John Evans.

Notes and references

- 1 P. Müller, *Cryst. Rev.*, 2009, **15**, 57–83; for an earlier view see P. G. Jones, *Chem. Soc. Rev.*, 1984, **13**, 157–172.
- 2 K. Biradha, C.-Y. Su and J. J. Vittal, *Cryst. Growth Des.*, 2011, **11**, 875–886; ed. L. Barbour, *Supramolecular Materials Chemistry in Supramolecular Chemistry: From Molecules to Nanomaterials*, eds P. A. Gale and J. W. Steed, John Wiley & Sons Ltd, Chichester, UK, 2012, pp 2790–3166.
- 3 J. Bernstein, *Cryst. Growth Des.*, 2011, **11**, 632–650.
- 4 A. Bingham, D. S. Hughes, M. B. Hursthouse, R. W. Lancaster, S. Tavener and T. L. Threlfall, *Chem. Commun.*, 2001, 603–604.
- 5 P. Coppens, *Angew. Chem., Int. Ed.*, 2009, **48**, 4280–4281.
- 6 F. P. A. Fabbiani and C. R. Pulham, *Chem. Soc. Rev.*, 2006, **35**, 932–942.
- 7 A. Dey, M. T. Kirchner, V. R. Vangala, G. R. Desiraju, R. Mondal and J. A. K. Howard, *J. Am. Chem. Soc.*, 2005, **127**, 10545–10559.
- 8 G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerriger, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Cryst.*, 2009, **B65**, 107–125.
- 9 See for example R. A. Jones, G. Wilkinson, A. M. R. Galas, M. B. Hursthouse and K. M. A. Malik, *J. Chem. Soc., Dalton Trans.*, 1980, 1771; E. W. Abel, M. M. Bhatti, M. B. Hursthouse,

K. M. A. Malik and M. A. Mazid, *J. Organomet. Chem.*, 1980, **197**, 345–355.

10 <http://www.epsrc.ac.uk/funding/facilities/epsrc/Pages/midrangeprog.aspx>.

11 <http://www.diamond.ac.uk/Home/Beamlines/I19.html>.

12 A. L. Fuller, L. A. S. Scott-Hayward, Y. Li, M. Bühl, A. M. Z. Slawin and J. D. Woollins, *J. Am. Chem. Soc.*, 2010, **132**, 5799–5802.

13 C. W. Lehmann, A. Karaulov and M. B. Hursthouse, *Zeitschrift für Naturforschung. A, A Journal of Physical Sciences*, 1993, **48**, 63–67.

14 D. Erdemir, A. Y. Lee and A. S. Myerson, *Acc. Chem. Res.*, 2009, **42**, 621–629.

15 For UK-based examples of funding agency policies see: <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>; <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/default.aspx>; <http://www.epsrc.ac.uk/about/infoaccess/Pages/roaccess.aspx>.

16 C. Nave and E. F. Garman, *J. Synchrotron Rad.*, 2005, **12**, 257–260; G. Evans, D. Axford and R. L. Owen, *Acta Cryst.*, 2011, **D67**, 261–270.

17 T. Schulz, K. Meindl, D. Leusser, D. Stern, J. Graf, C. Michaelsen, M. Ruf, G. M. Sheldrick and D. Stalke, *J. Appl. Crystallogr.*, 2009, **42**, 885–891.

18 B. Henrich, A. Bergamaschi, C. Broennimann, R. Dinapoli, E. F. Eikenberry, I. Johnson, M. Kobas, P. Kraft, A. Mozzanica and B. Schmitt, *Nucl. Instrum. Methods Phys. Res., Sect. A*, 2009, **607**, 247–249; P. Kraft, A. Bergamaschi, Ch. Broennimann, R. Dinapoli, E. F. Eikenberry, B. Henrich, I. Johnson, A. Mozzanica, C. M. Schlepütz, P. R. Willmott and B. Schmitt, *J. Synchrotron Rad.*, 2009, **16**, 368–375.

19 S. J. Coles and M. B. Hursthouse, *J. Appl. Crystallogr.*, 2004, **37**, 988–992.

20 B. B. He in *Two-Dimensional X-Ray Diffraction, Chapter 3. X-Ray Source and Optics*, Wiley, 2009, DOI: 10.1002/9780470502648.ch3

21 M. Otendal, T. Tuohimaa, U. Vogt and H. M. Hertz, *Rev. Sci. Instrum.*, 2008, **79**, 016102.

22 W. Clegg, A. J. Blake, J. M. Coles, J. S. O. Evans, P. Main, S. Parsons and D. J. Watkin, *Crystal Structure Analysis: Principles and Practice*, 2nd Edition, Oxford University Press, Oxford 2009.

23 A. J. Blake, W. Clegg, J. A. K. Howard, P. Main, S. Parsons, D. J. Watkin and C. Wilson, *Zeitschrift Fur Kristallographie*, 2002, **217**, 427–428.

24 <http://www.amercrystalassn.org/content/pages/main-summer-school-2010>.

25 <http://www.openresearchcomputation.com/>.

26 <http://www.ccp14.ac.uk/>.

27 <http://www.iucr.org/resources/commissions/crystallographic-computing>.

28 G. Oszlányi and A. Süto, *Acta Cryst.*, 2004, **A60**, 134–141; L. Palatinus and G. Chapuis, *J. Appl. Crystallogr.*, 2007, **40**, 786–790.

29 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339–341.

30 R. I. Cooper, A. L. Thompson and D. J. Watkin, *J. Appl. Crystallogr.*, 2010, **43**, 1100–1107.

31 G. M. Sheldrick, *Acta Cryst.*, 2008, **A64**, 112–122.

32 <http://www.cse.scitech.ac.uk/ccp/>.

33 T. Gelbrich and M. B. Hursthouse, *CrystEngComm*, 2005, **7**, 324–336; J. A. Chisholm and S. Motherwell, *J. Appl. Crystallogr.*, 2005, **38**, 228–231.

34 <http://www.ukoln.ac.uk/projects/ebank-uk/>; <http://www.ukoln.ac.uk/projects/L2S2/>; <http://research.microsoft.com/en-us/projects/orechem/>; C. Lagoze, *J. WebSci*, 2009, <http://journal.webscience.org/112/2/websci09-submission-10.pdf>; <http://www.mylabnotebook.ac.uk/>.

35 <http://www.jisc.ac.uk/publications/reports/2010/keepingresearchdatasafe2.aspx>; <http://www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf>.

36 S. Delagrière, P. Brenchereau, L. Launer, A. W. Ashton, R. Leal, S. Veyrier, J. Gabadinho, E. J. Gordon, S. D. Jones, K. E. Levik, S. M. McSweeney, S. Monaco, M. Nanao, D. Spruce, O. Svensson, M. A. Walsh and G. A. Leonard, *Bioinformatics*, 2011, **27**, 3186–3192.

37 For UK and US examples of a) Data Management and b) Open Access policies see: <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/default.aspx>; <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>; <http://www.epsrc.ac.uk/about/infoaccess/Pages/roaccess.aspx>; <http://www.rcuk.ac.uk/documents/documents/2006statement.pdf>.

38 S. Grazulis, D. Chateigner, R. T. Downs, A. T. Yokochi, M. Quirós, L. Lutterotti, E. Manakova, J. Butkus, P. Moeck and A. Le Bail, *J. Appl. Crystallogr.*, 2009, **42**, 726–729.

39 S. J. Coles, J. G. Frey, M. B. Hursthouse, M. E. Light, A. J. Milsted, L. A. Carr, D. DeRoure, C. J. Gutteridge, H. R. Mills, K. E. Meacham, M. Surridge, E. Lyon, R. Heery, M. Duke and M. Day, *J. Chem. Inf. Model.*, 2006, **46**, 1006–1016.

40 <http://www.reciprocalnet.org/index.html>.