

X-ray Crystallography. Is the Gold Standard becoming Tarnished?

The citation for the Fankuchen Award mentions my contribution to the teaching of crystallography, and my stewardship of the program CRYSTALS. Both of these fields of interest arose by chance rather than design. If I have been any good as a teacher, it was because I found the work of great crystallographers difficult to understand, and having worked to understand it, have tried to simplify it for other folk like myself. The success of CRYSTALS is entirely due to the many people who contributed to the infra-structure and the crystallographic utilities.

A Short Biography. By happy coincidence, I was invited to give the 2011 Fankuchen lecture in the year I was due to retire. This gave me the idea that the talk could be a kind of retrospective, a look back at how x-ray crystallography

had changed during the 47 years since I started work in 1964 with Tom Hamor, for my PhD in the Chemistry Department of Birmingham University. Tom was newly appointed that year, and I was his first student. The equipment he inherited was an old Leeds Weissenberg camera and a Phillips generator. Later he purchased a Stöe precession camera and an Enraf-Nonius Integrating Weissenberg. For my thesis I solved just three structures, two by visual estimation of intensities, and one using a Joyce-Loebl microdensitometer with the integrated films. At that time it would take months to collect multiple film packs for a 'full' two-axis set of Weissenberg data. Luckily I never had to use Beevers-Lipson strips because the university had recently acquired an English Electric Leo Marconi KDF9 digital computer. Tom and I were major users of the machine, running it over-night using a Fourier/isotropic least squares Autocode program written by John Rollett and his students in Oxford, or an Algol anisotropic refinement program written by Durward Cruickshank's group in Glasgow. At that time NMR and the optical spectroscopies were immature techniques, and x-ray crystallography was the Gold Standard for structure analysis.



Tom Hamor with Janet Cheatham, Univ. of Birmingham.

In 1967 I went to work in Oxford for one year as H. M. (Tiny) Powell's last post doc. My task was to get data out of the Hilger and Watts Royal Institution Y190 Linear Diffractometer. This machine was a mechanical implementation of the Ewald construction, and a great aid in making the reciprocal lattice very real. The crystal had to be aligned to the instrument axes, adjustments being made to the goniometer head with the shutter open. I still have all my fingers. Data collection times were directly proportional to the number of unique reflections (except for special zones, equivalents were almost never measured) and so could take weeks.

In 1967 I went to work in Oxford for one year as H. M. (Tiny) Powell's last post doc. My task was to get data out of the Hilger and Watts Royal Institution Y190 Linear Diffractometer. This machine was a mechanical implementation of the Ewald construction, and a great aid in making the reciprocal lattice very real. The crystal had to be aligned to the instrument axes, adjustments being made to the goniometer head with the shutter open. I still have all my fingers. Data collection times were directly proportional to the number of unique reflections (except for special zones, equivalents were almost never measured) and so could take weeks.



Keith Prout with Pete Symmonds ~2002, Oxford Univ.

When Tiny retired, in about 1972, Keith Prout became head of the Chemical Crystallography laboratory. Like Tom and Tiny, he was a very serious and knowledgeable crystallographer, with a real enthusiasm for encouraging and teaching. In fact, his passion for helping others led to his own career being neglected, and he never got the international recognition he deserved. The Linear H&W Y190 was later replaced by a Hilger and Watts Y290 four circle diffractometer, controlled by a PDP8 computer built from printed circuit boards with discrete transistors and a ferrite core memory. This, in its turn, was replaced by Enraf-Nonius CAD4 diffractometers, which brought data collection times down to a few days, and later, in 2000, by the truly brilliant Nonius Kappa CCD machines. The xCCD machines are still working side by side with dual microsource Oxford Diffraction (Agilent) instruments. In the 47 years since I took my first Weissenberg photographs, data collection times have dropped perhaps three orders of magnitude - from months to just hours, or even minutes.



David Watkin and J. O. W. R. Q. M., Oxford Univ.

Of course, during this same time computers and software also evolved. By 1967 Jeff Ford and John Rollett had replaced the old Oxford Autocode programs with a FORTRAN version. At that time every crystallographer had some programming skills so that Bob Carruthers, having come across a twinned crystal during his PhD

Of course, during this same time computers and software also evolved. By 1967 Jeff Ford and John Rollett had replaced the old Oxford Autocode programs with a FORTRAN version. At that time every crystallographer had some programming skills so that Bob Carruthers, having come across a twinned crystal during his PhD

work, was able to convert an existing program to deal with the data. Later, he went to work with Ricardo Spagna in Rome, helping write the code that later became Caos. Returning to Oxford, Bob and John began creating CRYSTALS, with the meticulous attention to detailed infra-structure that has enabled the program to evolve until the present day.

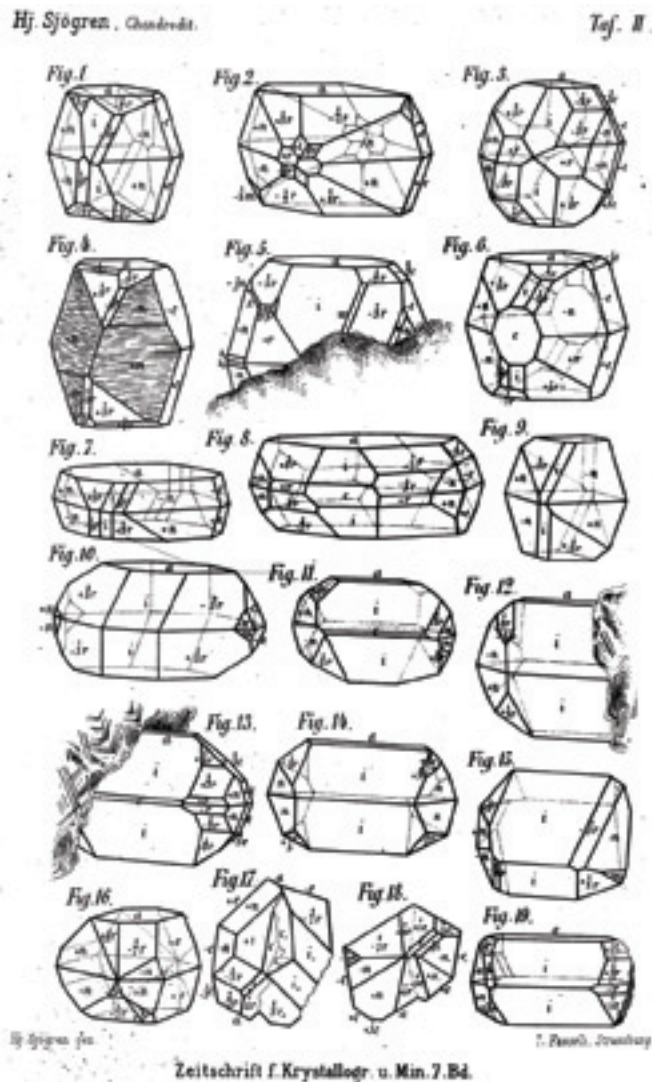
In about 1975 Bob announced that he was leaving Oxford to work with the super computer company Control Data Corporation, so that I found myself looking after not only the diffraction equipment, but also CRYSTALS. Keith Prout had always encouraged chemists to take a hands-on approach to crystallography, and visitors to the laboratory were often astonished to see project students using the instruments themselves. This endless stream of beginners led to my interest in teaching, and in trying to make our programs effective both as research and as teaching tools. During the VAX era, Stan Cameron sent Bev Vincent to work with us for a few months. Bev built our first *Interrogative User Interface*, which was later developed by Paul Betteridge into a sophisticated reprogrammable user interface. Also during this, time we tried to turn CRYSTALS into a fully automatic structure solution and refinement system. Then, as now, the automation ground to a halt whenever disorder turned up, and by 1995 Richard Cooper decided that it was probably easier to teach a chemist crystallography than to teach a computer chemistry, and he devised our first reprogrammable *Graphical User Interface*. Richard was much influenced by a close collaboration with Lachlan Cranswick, whose ideas for “eye candy” have now even found their way into programs like Olex2.

In 1976 Bob Sparks published a short program simulating refinement and useful for benchmarking computers in a crystallographic context, which was programmed into CRYSTALS. Sadly, we didn't record results from computers of that period, but we can compare a Microvax 3800 (1989) with a 3.0 Ghz Intel Duo (2010). We can now do in one day what would have taken three days in 2010, and almost five years in 1989. In high resolution (small molecule) crystallography, getting a reliable trial structure is no longer a serious issue. SHELXS soldiers on with the reliability of the Old Guard, new versions of SIRware bring new facilities, and Superflip has made charge flipping a standard technique. However, no really new ideas have emerged for working up structures, and we are still using methods which were readily available by the 1970s.

X-ray Crystal Structure Analysis in 2011. In 2011 we have diffractometers of great sensitivity, we have brilliant laboratory sources and we have immensely powerful synchrotron sources. Work at 100°K is routine, and at 10°K reasonably feasible. Diamond Anvil Cells enable samples to be examined at pressures up to 10 GPa. Computers work at breath-taking speeds.

In effect, we have amazing technology available. What science can we do with it?

"The age of intramolecular structural chemistry is declining for small molecules. There is very little that can be added to the average intramolecular geometrical data collected by use of the Cambridge Structural Database; anything at variance with these well-established averages is most probably wrong." (Gavezzotti & Flack, 2005, www.iucr.org/iucr-top/comm/cteach/pamphlets/21/21.html). Has small molecule crystallography become just stamp collecting?



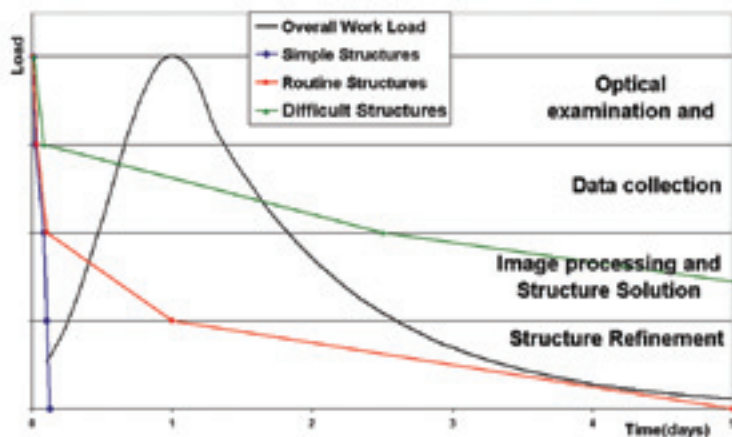
Indexing faces and measuring interfacial angles was legitimate research in Victorian times. Reproduced from Fig 2, Zeitschrift für Kristallographie, 1883.

Zeitschrift für Kristallographie, for the declining years of the 19th Century, contains page after page of beautiful drawing of crystalline minerals, and tables of the interfacial angles measured with improbable precision. With hindsight, we know that these measurements added nothing to our understanding of the Natural World, since the Law of Constancy of Interfacial Angles was already well established. Is *Acta Crystallographica E* just another repository for more irrelevant structures: the *Journal of Squeaky-Clean Structures*? Perhaps more interesting would be the *Journal of Rotten Structures*. To publish here authors would need a track record of ‘good’ structures to demonstrate their skills, but working up structures would be stubbornly intractable. The deposited material would include the author’s attempts, and all the diffraction data, including images. Keith Prout once said that there is no such thing as Bad Data, only Bad Models. This is perhaps an educational exaggeration (since a misaligned instrument will inevitably yield bad data) but the underlying thesis is that the solid state is probably more complicated than the analysts wish to admit. The truly troublesome structures are the ones which yield new insights.

On a one-by-one basis, “perfect” structures probably have a very limited crystallographic interest, but taken en masse they inform us about the physical world. At first sight the Cambridge Crystallographic Data Base might be seen as a grave yard for structures, but the software developed by the CDC turns data into knowledge, and the data base is now an arbiter of normality. It is well established that interatomic distances and bond lengths,

are influenced by the local environment. The program MOGUL enables an analyst to compare every bond in his current structure with those in the same environment in the data base. It provides a robust chemical yardstick by which to assess analyses, and should be made a mandatory part of the IUCr checkCIF process.

Structure Analyses - where next? Fully automatic structure analysis is just as far out of reach as it was in 1976, when Rollett thought it was just round the corner. Failure to correctly identify the type of every atom in the structure, and correctly insert or locate all hydrogen atoms means that the analysis has failed, and that the opinion of a human analyst is needed. Finishing anything but a trivial structure still requires a structure analyst to use his time and experience. The diagram below summarizes experience with the Analytical Ser-



Schematic time profiles for Simple, Routine and Difficult structures. The Poisson-like curve represents the numbers of structures brought to a publishable condition as a function of time taken for the analysis. Courtesy of Amber Thompson.

vice in Oxford. There are three bands, reading from the top. Initial evaluation of a sample requires human experience, but is relatively quick. The second band, data collection, may take from hours to days and is fairly predictable once initial diffraction images have been obtained. It is the bottom band, working up the data to a publishable conclusion, which is least predictable.

Direct methods and charge flipping are now very powerful procedures for phasing a Fourier map to the level that the atomic skeleton of the structure can be seen. However, physics alone cannot always make automatic unambiguous assignment of atomic sites. Worse than that, some molecules refuse to fit into normal space groups - they are disordered. Disorder is a bottomless

pit into which crystallographers throw time, effort and money. When disorder is eventually resolved, it is often clear that the resolution has added nothing significant to the scientific value of the structure. Was it worth the effort?

From time to time our administration asks us to provide a “Cost for a Structure Analysis”, expecting a simple answer. Of course, there is no simple answer. One way to evaluate the cost is:

$$COST = (\text{instrument cost}) / \text{write-off period} + \text{space charge} + \text{salary \& overheads} + \text{consumables}$$

The first three terms in this expression have to be paid even if the equipment is turned off so, in part, the cost of no structures is the same as the cost of some structures. The only saving which can be made by turning away a structure is that of the consumables – perhaps just \$30! Administrators don’t like this calculation, and so ask how many structures are done in a year (to try to get an average cost). Does “done” mean published, completed to a publishable standard and then left to rot on a hard disk, or completed just enough to confirm the identity, and, does it include work started but then abandoned for whatever reason? However ‘done’ is measured, a few things are possible to bring down the unit cost.

1. Put more samples on the instruments each day. This increases the burden on the analyst, and may even require additional analysts unless the data processing can be distributed over the end users, in much the same way that NMR data is.
2. Provide better automation for working up the data, and in particular for the treatment of disorder. It is a sad indictment that, in an age when

car registration plates are automatically read by computers at every garage, when OCR of old documents is routine and Immigration can automatically check passport holders against their photographs, we still have no software to automatically resolve structures. This will need a new approach to codifying chemical knowledge – it has been my experience that the best service crystallographers we have had in Oxford have also been first-class chemists.

3. Have a clear strategy to limit the amount of time spent on each structure.

Item 3 in this list is one which many analysts shy away from, yet in financially difficult times it is perhaps the most important. In the UK, research funding is not so generous that we can be wasteful of resources. Before starting an analysis it is important for the customer to declare what the analysis will be used for, for example if or how the work will be published. The

experiment and data processing can then be tailored to those aims. For example, don’t collect 10 equivalent reflections when just a few will do. Occasionally an analysis proves to be much more important than was foreseen. In general, re-measuring a few data sets is a better use of resources than collecting everything to the highest standards first time around. Joe Reibenspies, in an earlier session of the 2011 ACA meeting, declared himself unhappy with this approach, saying that while he was working on a structure it was “his”, but when he published it, it was “ours”. A seductive argument, but I believe that when I do work, it is for the Chancellor, Masters and Scholars of the University of Oxford, who may agree to share their results through publication. The old mineralogists measured interfacial angles with great care, ‘just in case’. You can be reasonably sure that this accuracy was rarely needed, and if careful measurements were needed, the material was reexamined.

The Gold Standard.

Few people regularly eat from a golden platter; we don't measure the timber for a chicken hut with a micrometer. Crystallography can still be the Gold Standard, but a gold standard is not needed for all work. The analyst simply needs to demonstrate fitness-for-purpose. If the customer changes the purpose later, that is a separate issue. The problem for journals is to understand the purpose of the work, and see if it is satisfied. The purpose cannot be evident from the information currently included in a CIF, and it is therefore unlikely that a program can make this decision. However, checkCIF may help the referee decide if the analysis is 'good enough' for the purpose described in the text. In the future, CIFs may also need to record what the analysis was 'good enough' for. If someone has a different purpose, they may need to re-measure the data.

Small Molecule Crystallography – Where Next?

For over half a century small molecule crystallography has been driven (and funded) by the chemist's interest in molecular structure. Chemistry can still be crystallographically challenging, for example the study of excited states. However, for the professional crystallographer, the molecular solid state has become interesting in its own right. We understand quite a lot about strong interatomic interactions - those which form 'bonds', but very little is known about the weaker interactions - the ones which enable molecules to form solids. Structure prediction is scarcely feasible, morphology prediction even less reliable, and we can only rarely predict the conditions needed to grow decent crystals of a novel material. If a proper amount of effort were expended in automating the final stages of analyses, we could say that structure analysis was 'Done and Dusted', and crystallographers could turn their attention away from determining molecular geometry, and towards developing an understanding of the solid state. This is a period of renaissance for crystallography, with new horizons wide open for adventurous young scientists.



David Watkin

Cartoon courtesy of Nick D. Kim. See page 10.

What Is the Future for Crystallography?



With all the hoopla surrounding the primary and national elections, are there any things we can be reasonably sure of regarding the prospects for crystallographers? I hope so and I think so. Whenever you start talking about where the US needs to make progress in education it always comes down to STEM (science, technology, engineering, math). Crystallographers certainly exercise their skills in all of these areas and have graduate students adept in one or another of these disciplines. Admittedly we no longer lead the way in usage of 3D graphics, having yielded that position long ago to the "gamers", but we continue to push useful graphical applications. We use many different display options and have tools for figure creation in applications, e.g. docking simulations based on multiple structures. Computer coding, whether it's for structure solution and refinement packages, robotic control software, or edge detection for crystallization imaging is based on fundamental mathematics. Other aspects of what we do are founded on practicing science, from biology through chemistry and physics to material science and geology. Our field relies on and enhances STEM.

While NASA may shudder at questions about the relevance of the space program in these tight fiscal times, crystallography remains the primary means of determining the atomic structure of materials all the way from small molecules to viruses and structural complexes. And the solutions of these structures are highly relevant to the real problems that we face today. From the structure of an iridium bromide complex with potential for hydrogen storage to elucidating the mechanism of plant response to stress as well as the structures with relevance to the biomedical fields, *crystallography really matters*. Our ACA SIGs range from highly applied to highly theoretical, from problem solving to problem proposing, but all share in their dependence on density from Fourier transforms of diffraction. Wouldn't the early pioneers of crystallography be amazed by how far we have come, not to mention how diverse we are?

As a university professor who teaches experimental physics, I often wonder about the new generation of students coming through my classes. They don't even know how to use an Abstract Index (I haven't seen one in awhile, now that I think about it.) They're much less likely to know constants, remember formulae, or even, in some cases, own their own textbooks. In 10 more years their bookshelves will likely be filled with Ipads, Kindles, Nooks or simply drive space in the "cloud" (for which I have little trust). The things they excel at are searching, finding, and concatenating. It surprises them that I know integration techniques other than Maple or Mathematica and sometimes get a simple answer where theirs are hopelessly complex. But nevertheless, the imagination and creativity of this generation, the perseverance when interest is captured and the curiosity for unknown things still runs strong. Our students may come from farther away but they will continue to advance the field, innovate, and produce both our software and our hardware.

So, looking forward, whether the blue states overwhelm the red or vice-versa, I'm sure there will be difficulty with funding. And there may be scarcity of resources as more is asked for and less provided. However, I am equally certain that our field will continue to advance, engage new minds, and provide a basis for solving many of the truly important, non-political problems that arise.

Ross Reynolds

