Automated Chemical Crystallography

Amy L Fuller, Lindesay A. S. Scott-Hayward, Yang Li, Michael Bühl , Alexandra M. Z. Slawin* and J. Derek Woollins*

School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9ST (UK)

RECEIVED DATE (automatically inserted by publisher); E-mail: jdw3@st-andrews.ac.uk

Abstract: The first fully automated small-molecule robotic X-ray diffractometer is described. After demonstrating the utility of the instrument using multiple samples of ammonium bitartrate we investigated the conformational chirality of diphenyl dichalcogenide (E_2Ph_2, where E = S, Se, or Te). Structural and computational studies suggest that the two enantiomers are energetically indistinguishable. Therefore, it was unsurprising that we found (in 35 suitable data collections) the proportion 0.51:0.49 of *M*-S_2Ph_2 to *P*-S_2Ph_2 in the bulk sample. Interestingly, after 65 data collections of Te_2Ph_2, (46 provided suitable data sets), we found the proportion 0.72 $\pm$ 0.13 of *M*-Te_2Ph_2, Suggesting there could be a statistically significant preference for the *M*-enantiomer in the sample examined here. We found that Se_2Ph_2 underwent homochiral crystallization with all 24 crystals being *M*. Our experiments may represent a salutary lesson in statistical analysis.

Introduction

Completely automated chemical crystallography has the potential to change the way crystallography is viewed. Knowledge of molecular shape is a key component in the understanding of molecular properties and reactivity. Chemists use mainly NMR and X-ray crystallography to determine the topology and stereochemistry of synthetic molecules. Recently, there have been rapid advances in these methods enabling improvements in the speed and precision. The most accurate and complete information comes from X-ray crystallography, which allows full 3-D information to be obtained in a matter of hours[1,2] It is interesting to compare the progress in NMR spectroscopy with that of X-ray diffraction. In NMR, there was a step function improvement with the advent of FT techniques and superconducting magnets, which led to the ability to measure spectra rapidly and within a decade, the use of autosamplers became routine in many laboratories. In crystallography, the use of CCD detectors (covering an area in space and allowing many reflections to be measured simultaneously) has had a dramatic effect on the rate at which data can be collected but has not been matched by improvements in automation. Here we describe our progress in developing an automated system for chemical crystallography. The use of a robotic arm to exchange samples, along with automated software enables full utilization of the data collection capability of a modern CCD laboratory based X-ray diffractometer.

Results and Discussion

In designing a fully automated X-ray diffractometer capable of continuous operation our ultimate aim is to provide a system, which after submission of a crystalline sample; collects data images, indexes the observed reflections to give a unit cell, optimizes the data collection strategy, integrates the observed data to give a reflection (hkl) file, solves and refines the structure and finally informs the sample submitter of the outcome. Clearly approaching this level of automation for a large proportion of samples is very demanding, and requires considerable development beyond the sample changing automation available in macromolecular crystallography, but we have now accomplished this to a significant degree of success.

The diffractometer was constructed using a conventional CCD detector and goniometer with the addition of an electronic goniometer head (to allow for automatic centring of crystals) a robotic arm and two drawers each capable of holding 48 crystals mounted on magnetic pins. To enable sample mounting by inexperienced users we chose to mount crystals on loops. The user(s) mount their crystals on loops with magnetic bases[4] and places them in one of the four sample trays. After this point the process is fully automated right through to structure solution if requested. Using a robotic arm, (which was originally developed for protein studies) the crystal is first transferred to the goniometer head and is then automatically centred. Crystals in chemical crystallography are usually grown with a diverse range of shapes, sizes and colors. Previously available software for centring large, homogeneous and usually colorless protein crystals was not adequate for small molecule crystal centring.

The development of new image recognition and centring software has enabled good crystal centring from a wide range of samples types. Successful centring is achieved in >90% of cases, some examples are shown in Figure 1. The established diffractometer control software was developed so after trial images have been taken, the unit cell is determined and the crystal quality ranked. The user can define a minimum quality so that if the ranking fails, it moves on to another crystal. However, if the rank passes then the data are collected and integrated and subsequently the structure is determined using system STARRY[3] without human intervention. System STARRY has been developed to run automatically all of the processes from Direct Methods solution

through to refinement, utilising relevant standalone crystallographic programs (currently programs from the SHELX family are used) The design philosophy is to imitate the crystallographers' flexible thinking during structure development by using the indications from the refinement and structural knowledge, rather than relying on the proposed exact chemical formula. Finally, after structure solution and refinement, an email alert is sent to the user informing them of the data collection status. As a fallback, the data are also available for the users to return to if intervention is required eg especially for a particularly difficult sample. The system has been developed sufficiently that it is now available commercially[4]

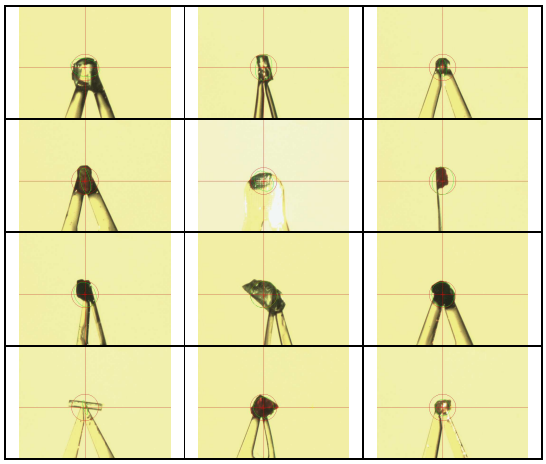

Figure 1. Examples of the outcome of automatic crystal centring using different size/shape crystals. The red circle represents the central portion of the X-ray beam.

As one test of the system we mounted 24 different sized crystals of ammonium bitartrate and collected data automatically. Of these 24 samples, 19 gave successful structure determinations without any user intervention (Table 1). The results indicate good quality results from a range of crystals with fair *R*int and R_1 values. values . We have further tested our system with >1500 samples ranging from organic molecules through coordination chemistry compounds to 'inorganic solids', such as coordination polymers and aluminium phosphates.

Typically we obtain ca. 80% success in acquiring useful hkl files from the fully automated sample collections, with a somewhat lower success rate for full automated structure determination. In continuous operation, for 'normal' CCD quality/sized crystals, we are able to collect and process up to 20 full datasets per 24 hours.

This increase in throughput is leading to a different philosophy in how we use X-ray crystallography. Since we can now quite easily obtain 6-10 structures overnight, it is possible to use this equipment as an analytical as well as a structural tool. Furthermore, the ability to collect several datasets means we can measure several crystals of a particular compound, then rank them competitively before collecting the best data set, or automatically test for polymorphism and/or test enough crystals to ensure that the chosen crystal is 'typical' This ultimately improves data quality with very efficient use of both operator and diffractometer time.

ID	Rank	R*int*	Total no. Refl.	No. of unique	No. of Obs	R_1
01	218	0.059	2375	1290	1227	0.065
04	214	0.061	3012	1310	1288	0.053
05	191	0.041	2252	1269	1245	0.057
06	216	0.167	2228	1272	1162	0.125
07	214	0.058	2558	1263	1193	0.09
09	217	0.033	2210	1228	1183	0.058
10	198	0.14	2152	1286	1259	0.089
12	189	0.11	4051	1308	1279	0.067
14	201	0.024	2312	1239	1217	0.048
15	211	0.041	2587	1271	1190	0.066
16	209	0.047	2383	1287	1215	0.063
17	203	0.041	2799	1271	1254	0.053
18	215	0.050	3092	1245	1212	0.056
19	205	0.036	2585	1279	1249	0.054
20	195	0.085	1808	1143	1122	0.069
21	186	0.038	2267	1266	1242	0.050
22	222	0.046	2198	1274	1250	0.057
23	201	0.078	2258	1267	1215	0.066
24	209	0.078	2415	1300	1214	0.075

Table 1 Selected parameters from a study of ammonium bitartrate crystals, $[NH_4][C_4H_5O_6]$ All data were collected at 125 K. The molecule crystallises in the orthorhombic space group P212121. Typical standard uncertainties of cell parameters 0.002-0.003 Å. Values are for data as automatically processed, agreement factors are for a model with no hydrogen atoms, all non-hydrogen atoms refined with anisotropic thermal parameters. Supplementary Table s1 contains details of crystal dimensions.

One common occurrence in single crystal studies is the determination of a structure with a particular absolute configuration from a single crystal without any evidence that this crystal is 'typical'. To further test the abilities of our robotic system we investigated the usefulness of single crystal studies for the determination of the crystallographic enantiomorphic excess in PhEEPh (E = S, Se or Te). There was some low resolution data available on these compounds[5] but as a starting point we crystallographically characterized both *M*- and *P*-enantiomers of each compound in the E_2Ph_2 series (Supplementary Tables 2, 3 Figure 2). The six structures are isomorphous. Unsurprisingly, the E─E bond lengths increase as the chalcogen size increases, from S (2.0260(14) Å, 2.0289(7) Å) to Se (2.3066(7) Å, 2.3073(10) Å) to Te (2.7089(7) Å, 2.7073(5) Å) and are statistically indistinguishable between the *P*- and *M*- enantiomers of the same chalcogen. Logically enough, as the chalcogen atom

size increases, the E─C bond distance also increases and again, a negligible difference is seen between the enantiomers.

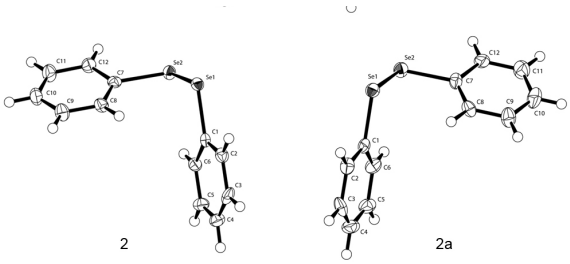

Figure 2. X-ray structures of *P*-Se_2Ph_2 (**2**), *M*-Se_2Ph_2 (**2a**) The S and Te analogues are isomorphous.

The E-E-C angles decrease across the series from S to Se to Te, however there is a slight difference between the E(2)-E(1)-C(1) angle, which in every molecule is ~1° to 3° larger than the E(1)-E(2)-C(7) angle. In S_2Ph_2 and Se_2Ph_2, the similar E(1)-C(1)-C(2) and E(2)-C(7)-C(12) are much smaller than the E(1)-C(1)-C(6) and E(2)-C(7)-C(8) angles. In Te_2Ph_2, all four of these angles are similar. Finally, the torsion angles C(1)-E(1)-E(2)-C(7) in both *P*- and *M*-Te_2Ph_2 are 90.7(3)° and -90.5(2)°, respectively. The same torsion angles are slightly smaller (~ 5°)in the sulfur and selenium analogues.. With the knowledge that the individual enantiomers are structurally indistinguishable, other than the direction of their helicity, we investigated the energetic differences between the conformations. Rotational barrier calculations on free Se_2Ph_2 resulted in one minimum where the molecule has C_2 symmetry (the *M*- or *P*-enantiomer), as well as two transition states for rotation, with *syn*- and *anti*- conformations of the two phenyl groups (Figure 3). Both DFT and MP2 calculations predict a barrier on the order of 5-6 kcal/mol for the lower of the two transition states. This value is.consistent with a previously calculated rotational barrier of 12 kcal/mol (at 204 K) for a bis(2,4,6-tri-tert-butylphenyl) diselenium derivative.[6] For Te_2Ph_2, a smaller rotational barrier than Se_2Ph_2 is expected, and the conversion barrier in the bis(2,4,6-tri-tert-butylphenyl) ditellurium derivative is 9.4 kcal/mol.[7]

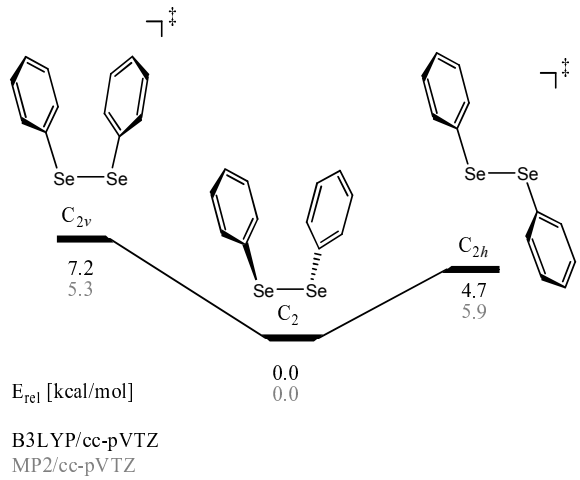

Figure 3. Results from Se_2Ph_2 rotational barrier calculations (B3LYP/962(d)/6-31G* geometries employed, see Supporting Information for details).

As one would expect in the absence of a chiral environment, the *M*- and *P*-forms of the C_2 minima are isoenergetic[9] and our structural and computational studies confirm that these conformational enantiomers are indistinguishable. However previous data (although limited) suggested that there could be a preferential crystallization processes possible for these systems.[8] This phenomenom is not unknown and Kondepudi *et al* [9] studied (using optical examination of the products) the crystallization of aqueous sodium chlorate solutions which gives D and L forms of solid $NaClO_3$ and noted that autocatalysis/nucleation plays an important role in determining the outcome. Stirred solutions underwent spontaneous resolution, having a clear preference for one hand or the other whilst static solutions gave mixtures of D and L enantiomers. To produce statistically significant data in our experiments, we needed a large number of randomly chosen crystals. Crystals of commercially available samples were used, but we have not examined multiple batches. We used a simple Flack parameter test[10] as the criterion for handedness though we recognize that there is the possibility of solid solutions causing some errors in this approach.

Forty-nine single crystals of S_2Ph_2 were analyzed by X-ray crystallography. After data collection and integration, the structures were solved using direct methods, refined anisotropically and the Flack parameter was closely examined. Of the 49 samples, 35 had Flack parameters and *R* factors of an appropriate quality to be reasonable for the determination of chirality. An analysis of these 35 samples showed that 18 crystals were the *M*-enantiomer.

The estimated proportion of *M*-S_2Ph_2 in the sample bottle, $\hat{p}$, can be found by using Equation 1, where M is the total number of *M*-enantiomers and X is the total number of experiments. The estimated proportion of *M*-S_2Ph_2 in our sample bottle is 0.51. This near 50:50 ratio of *M*-:*P*- is consistent with the structural and computational data.

$$\hat{p} = \frac{M}{X} \qquad \text{Equation 1}$$

Unfortunately, Se_2Ph_2 was delivered as a powder. Recrystallization from CH_2Cl_2, yielded yellow crystals, of which 24 single crystals were analyzed by X-ray diffraction. Out of the 24 crystals, 17 produced suitable data sets, all of these solved as the *M*-enantiomer. It appears that this sample, like those of Shimizu *et al.*,[14] underwent homochiral crystallization.

Finally, sixty-five single crystals of Te_2Ph_2 were analyzed.. Out of 46 good quality data sets, 33 were *M*-Te_2Ph_2. Using Equation 1, the estimated proportion of *M*-enantiomers in our sample bottle is 0.72.

This result suggests that in the sample analysed here there is an enantiomeric preference for *M*-Te_2Ph_2, but the error in the experiment needs to be found. The error associated with the estimate, $\hat{p}$, is found through calculating a 95% confidence interval (CI_{95}) which gives a range within which we are 95% certain that the true proportion of *M*-E_2Ph_2 in the sample bottle will lie. Since we have proportion data, a binomial distribution is used to calculate CI_{95} (Equation 2). A binomial distribution usually applies when an experiment is repeated a fixed number of times. Each trial has one of two outcomes - success or failure, or in this case - *M*-E_2Ph_2 or not. The probability of success is the same for each trial and the trials are statistically independent of each other. A summary of the results for the confidence intervals is shown in Table 1.

$$CI_{95} = \hat{p} \pm 1.96 * \sqrt{\frac{\hat{p}(1-\hat{p})}{X}}$$

Equation 2

The CI_{95} for *M*-Te_2Ph_2 is 0.72 ± 0.13, suggesting that there is a 95% chance that the true proportion of *M*-Te_2Ph_2 in our sample bottle is within the interval 0.59 – 0.85. Most remarkably, we appear able to crystallize the selenium compound entirely as one isomer.

The CI_{95} for *M*-S_2Ph_2 is 0.51 ± 0.17. These results, as we expected from our structural and computational studies, suggest that in the sample bottle, crystals of S_2Ph_2 crystallize indiscriminately as either the *M*- or the *P*- enantiomer. This agrees with the expectations that there would be no noticeable preference for either enantiomer.

Table 2. Summary of CI_{95} for E_2Ph_2 experiment

	Total	Suitable	%*P*-	%*M*-	CI_{95}
S_2Ph_2	49	35	0.49	0.51	± 0.17
Se_2Ph_2[a]	24	17	0	100	
Te_2Ph_2	65	46	0.28	0.72	± 0.13

It is clear from the above that reported 'chiral'crystal structures must be treated with some caution unless supported by other data. The crystallization of only one isomer of Te_2Ph_2 is reminiscent of the experiments of Kondepudi and suggests autocatlaysis/nucleation. These experiments highlight an important pedagological issue with reproducibility in structural science and we intend expending further effort to assess how many duplicate measurements are needed in any particular situation.

In conclusion, using the **St And**rews **A**utomated **R**obotic **D**iffractometer (**STANDARD**), the crystals can be mounted, images collected and structures solved minimizing not only the time of the experiment, but the time required of a crystallographer. The developments we report suggest that single crystal studies could become even more accessible enabling higher quality data, multiple sampling for phase or chirality screening as well as being a valuable analytical tool.

Experimental Section

Samples came from commercial sources, Aldrich Chem Co or Alfa Aesar.

The automated diffractometer[4] consists of a conventional sealed tube X-ray source (Mo radiation) equipped with graphite monochromator, Rigaku Saturn 724 CCD, AFC10 goniometer, MicoglideTM three axis goniometer head, ACTOR™ Robotic arm and X-Stream LT device. Control software *Cameraman*[11] centres the crystal whilst *Director*[12] provides input and inventory control over the samples. Two independent drawers fitted with two removal trays which each hold up to 24 samples means that the maximum capacity of the queueing system is 96 samples. The independent drawers can be loaded seperately thus enabling continuous operation. *Director* automatically starts *CrystalClear*[13] to control data collection including indexing. If indexing results in a reasonable quality 'ranking' then the system goes on to collect a hemisphere of data either based on the indexed cell using a calculated strategy or using a standard data collection which is set up in the recipes. One the data collection is underway a second PC automatically integrates, Laue checks, absorption corrects and outputs an hkl file. If instructed to the second PC will also attempt a fully automated structure solution[14, 15] At the end of these processes the sample submitter is emailed indicating the overall outcome.

To allow for the robotic mounting magnetic bases are used. We used conventional nylon loops or laser cut loops mounted on stainless steel pins from Mitegen. http://mitegen.com/mic_catalog.php?c=MicroMounts http://www.jenabioscience.com/cms/en/1/catalog/731_micromounts8482_19_mm.html

Experimental details for the E2Ph2 samples are included in the supporting information

Acknowledgment. We are grateful to the UK Engineering and Physical Sciences Research Council for support and to Russ Athay and Claire Wilson of Rigaku Americas for software development and support. M.B. wishes to thank EaStCHEM for support and access to the EaStCHEM Research Computing Facility maintained by Dr. H. Früchtl.

Supporting Information Available: Table 1 Crystals sizes and R factors of automated data collections for ammonium bitartrate. Table 2 Details data collections for 1-3, Table 3 selected bond lengths and angles for 1-3 and X-ray diffraction (CIF) details. This material is available free of charge via the Internet at http://pubs.acs.org/.

(1) W Massa, Crystal Structure Determination. 2004, Berlin: Springer. ISBN 3540206442.
(2) For a series of articles illustrating the power of X-Ray diffraction see: Foundations of Crystallography, 60 years of Acta Crystallographica and the IUCr, Acta Cryst A, 2008 64, Part 1 1-271
(3) Li, Y; Slawin, A.M.Z., St Andrews, Rigaku, Robotic by Yang (STARRY) uses a new set of algorithims loosely based on System Y which was briefly reported in M. B. Hursthouse, M.B.; Light, M.; Li,Y. Acta Cryst. 2004. A60, s81.
(4) For a video of the system described here in operation see http://www.rigaku.com/automation/actor-sm.html which illustrates the design described here
(5) PhSSPh: Lee, J.D.; Bryant, M.W.R., *Acta Cryst*. **1969** *B25* 2094-2101.; PhSeSePh: Marsh, R.E., *Acta Cryst*. **1952** *5*, 458; PhTeTePh: Llabres, G.; Dideberg, O.; Dupont, L. *Acta Cryst*. **1972**, *B28* 2438
(6) Mario Pinto, B.; Leung,R.N.Y.; Sharma, D. *Magn. Reson. Chem*. **1988**, *26*, 729-734.
(7) Due to parity violation of the weak nuclear force, the respective enantiomers are not exactly isoenergetic, but should differ slightly in energy; for H2Se2 and H2Te2, for instance, this energy difference has been predicted to be on the order of 10-12 to 10-11 kcal/mol, i.e., indetectably small, see e.g.Berger, R.; van Wullen, C. *J. Chem. Phys*. **2005**, *122*, 134316-8.
(8) Shimizu, T.; Isono, H.; Yasui, M.; Iwasaki, F.; Kamigata, N. *Org. Lett*. **2001**, *3*, 3639-3641.
(9) Kondepudi, D.K.;.Kaufman , R.J.; Singh N. *Science*, **1990**, *250*, 975-976.
(10) Flack, H.D.., *Chirality* **2008**, *20*, 681-690.
(11) Rigaku Corp. (2005). *Cameraman*. Rigaku Corporation, 3-9-12 Akishima, Tokyo, Japan.
(12) Rigaku Corp. (2009). *Director* Version 3.0. Rigaku Corporation, 3-9-12 Akishima, Tokyo, Japan.
(13) Rigaku Corp. (2009). *CrystalClear*. Version 2.0. Rigaku Corporation, 3-9-12 Akishima, Tokyo, Japan.
(14) Rigaku (2006). *Crystal Structure*, Version 3.8. Single Crystal Structure Analysis Software. Rigaku/MSC, 9009 TX, USA 77381-5209. Rigaku, Tokyo196-8666, Japan.
(15) Sheldrick, G. M.. Acta Cryst. **2008**, A64, 112–122.

Robotic, automated single crystals studies could become as convenient as autosamplers in NMR; this new system was used to assess enantiomeric excess in PhEEPh.