

Polymorph Identification and Crystal Structure Determination using a Combined Crystal Structure Prediction and Transmission Electron Microscopy Approach

Mark D. Eddleston,^[a] Katarzyna E. Hejczyk,^[a] Erica G. Bithell,^[b] Graeme M. Day*^[a+c] and William Jones*^[a]

Abstract: Electron diffraction offers advantages over X-ray based methods for crystal structure determination as it can be applied to sub-micron sized crystallites, and picogram quantities of material. With molecular organic species, however, crystal structure determination with electron diffraction is hindered by rapid crystal deterioration in the electron beam, limiting the amount of diffraction data that can be collected, and by the effect

of dynamical scattering on reflection intensities. While automated electron diffraction tomography provides one possible solution, in this paper we demonstrate an alternative approach where a set of putative crystal structures of the compound of interest is generated using crystal structure prediction methods, and electron diffraction is used to determine which of these putative structures is in agreement with the available electron

diffraction data. This approach enables the advantages of electron diffraction to be exploited, while avoiding the need to obtain large amounts of diffraction data or accurate reflection intensities. We demonstrate the methodology using the pharmaceutical compounds paracetamol, *scyllo*-inositol and theophylline.

Keywords: Structure elucidation • Analytical methods • Computational chemistry

Introduction

The ability to routinely solve crystal structures by X-ray diffraction was a key scientific achievement of the 20th century,^[1] and is a cornerstone of analytical chemistry. There are, however, limitations with the technique, especially for weakly diffracting organic compounds where it is not always possible to grow crystals of suitable size and quality for structure determination, even when using a synchrotron X-ray source.^[2, 3] More recently, crystal structure determination using powder X-ray data through direct methods or Rietveld analysis has become a routine process,^[3, 4] but is still unlikely to be successful with multi component mixtures, or with binary mixtures where the crystalline form of interest is present as a minor phase.

For beam stable materials, alternative approaches to crystal structure determination using transmission electron microscopy (TEM) have been developed over recent years.^[5, 6] These approaches are especially powerful for the characterisation of minor phases in mixtures: provided that a suitable crystallite can be located, the analysis can be performed using exceptionally small amounts of material (< 1 fg).^[7] The particular value of TEM lies in the possibility which it offers for combining imaging, diffraction and spectroscopic data from the same sample, with the potential for simultaneous acquisition of different data types. For example, in scanning mode (STEM) direct imaging of the crystal lattice can be coupled with spectroscopic analysis to yield chemical information at the unit cell level.^[8] In imaging mode the technique is routinely used for identification and characterisation of defect structures,^[9] in diffraction mode, detailed investigations of both symmetry^[10] and crystal structure^[11] are possible. Precision control of sample position and tilt, coupled with rapid acquisition of digital images, has contributed strongly to the development of holographic methods for the measurement of local fields within structures,^[12] and tomographic approaches to three dimensional characterisation of microstructure.^[13] Studies of inorganic materials have tended to exploit the high spatial resolution of the technique, which is now better than 0.1 nm for high performance instruments.^[14]

The above approaches are not, however, routinely applicable to materials such as organic molecular compounds, which are particularly sensitive to electron beam damage. In such systems crystal structure determination by electron diffraction is rarely achieved as the lifetime of even the most stable organic sample in an electron beam tends to be at most a few minutes,^[15] which is usually

[a] Dr M. D. Eddleston, Dr K. E. Hejczyk and Prof W. Jones
Department of Chemistry
University of Cambridge
Lensfield Road, Cambridge, CB2 1EW, United Kingdom
Fax: (+44) 1223 762829
E-mail: wj10@cam.ac.uk

[b] Dr E. G. Bithell
Department of Materials Science and Metallurgy
University of Cambridge
Pembroke Street, Cambridge, CB2 3QZ, United Kingdom

[c] Dr G. M. Day
Department of Chemistry
University of Southampton
Southampton, SO17 1BJ, United Kingdom
E-mail: g.m.day@soton.ac.uk

less time than would be required to collect a sufficient number of diffraction patterns for a full structural analysis. For the same reason, direct imaging of the crystal lattice is not usually possible due to the high beam fluxes required when working at high magnifications.^[16] In addition, space group determination is often hindered by the combination of anisotropic crystal growth morphologies and the limited tilt range in a TEM microscope, meaning that diffraction from certain crystal planes is rarely, if ever, observed. Additionally, as a result of multiple (dynamical) scattering it is difficult to obtain kinematical diffraction intensities, even with specialised techniques such as precession electron diffraction,^[17] and reflections are frequently observed where systematic absences would be expected.^[5, 18] Approaches to crystal structure determination based on automated electron diffraction tomography are currently being developed in order to overcome the difficulties described above.^[19]

Crystal structure prediction (CSP) is a computational approach to investigate the range of low energy crystal structures available to a compound. The most commonly used approach is global lattice energy minimisation,^[20] where putative crystal structures are generated by computationally exploring the crystal packing variables (unit cell dimensions, positions and orientations of molecules in the asymmetric unit), often in a restricted set of the most commonly observed space groups. These structures are energy minimised and ranked by lattice energy, with the assumption that the lowest energy structures are most likely to be observed. One of the main uses of CSP calculations is to assess the likelihood of finding new polymorphs of a compound. For example, the calculations might generate structures that have not been observed experimentally, but that are predicted to have similar, or lower, energies than the observed crystal form of the compound. Frequently, the outcome has been an indication of hitherto unknown polymorphs which have subsequently been obtained by further experimentation.^[21] CSP has also been successfully used to aid in the determination of crystal structures when it is not possible to grow crystals large enough for structure solution by single crystal X-ray diffraction and only powder data is available.^[22, 23] Typically, experimental powder X-ray diffraction (PXRD) patterns are compared with simulated powder diffraction patterns calculated for each of the low energy computationally generated crystal structures to identify which of these structures matches the experimentally observed crystal form. Recently, a similar approach using solid-state NMR (SSNMR) has been reported,^[24] where chemical shifts calculated from predicted structures are matched against experimental spectra.

In this paper, we report characterisation of the pharmaceutical compounds paracetamol, *scyllo*-inositol and theophylline by TEM to demonstrate the use of electron diffraction for the identification of the polymorphic form of individual crystallites in samples via a process of indexing diffraction patterns to one of the known crystal structures of the appropriate compound. The question then arises as to how the diffraction patterns might be interpreted if the crystal structure of the phase being analysed has not been previously determined. To answer this question, we consider the results of CSP, where sets of putative crystal structures for paracetamol, *scyllo*-inositol and theophylline were generated '*in silico*', and assess which, if any, of these potential structures permit satisfactory indexing of the observed electron diffraction patterns. Using this approach, we demonstrate that, had the crystal structures of the pharmaceutical phases investigated in this study not been determined previously, it would still have been possible to identify a

plausible crystal structure for these phases. The key advantage of this methodology for crystal structure determination is that a relatively small number of electron diffraction patterns are required (limited only by the need to establish agreement to only one of the putative structures), and these can be collected before a significant amount of electron beam damage has occurred in the sample. In addition, the process of using reflection intensities to locate atoms and molecules within a unit cell is avoided, though intensities can still be tactically used as a further guide to establishing that a correct crystal structure has been identified.

In a second paper,^[25] we highlight a scenario where during post-analysis it is determined that electron diffraction patterns from a crystal cannot be indexed to a known crystal structure suggesting that a new crystal form may have been observed, but where it is not possible to go back to the specimen for further analysis (for example a metastable crystal form which cannot be isolated reproducibly and which rapidly converts to a more stable phase). We demonstrate that the combined electron diffraction and crystal structure prediction approach to crystal structure determination is applicable in such a scenario, using as an example the identification of a previously unknown polymorph of the compound theophylline which has only been observed once during TEM studies with this compound. This analysis was performed on a crystal with a sub-micron thickness, a mass of approximately 3 pg, and despite the new polymorph existing as a minor component in a mixture with Form II of theophylline at a concentration below the limit of detection of analytical techniques such as powder X-ray diffraction. We also discuss the complementary aspects of approaches to crystal structure determination based on electron diffraction / crystal structure prediction and based on automated electron diffraction tomography.

Results

Phase identification with Form I of paracetamol

A TEM image and electron diffraction pattern recorded from a crystal of paracetamol prepared by crystallisation from the melt are shown in Figure 1. It was possible to perform this analysis before a significant amount of beam damage was induced in the crystal. The electron diffraction pattern shows a regular array of reflections indicating that the electron beam was aligned close to a zone axis of the crystal when the pattern was recorded. The average distances between reflections in two orthogonal directions were measured and used to calculate the d-spacings and interplanar angle of two low index crystal planes that were suitably aligned with the electron beam for diffraction to occur when the pattern was recorded. Using these values ($d_1 = 11.51 \text{ \AA}$, $d_2 = 5.65 \text{ \AA}$ and $\theta = 85.4^\circ$), an attempt was made to index against the reported crystal structures of paracetamol, Forms I, II and III (Cambridge Structural Database (CSD) structures HXACAN01, HXACAN23 and HXACAN29 respectively)^[26-28]. The pattern was found to be consistent exclusively with diffraction from the $\langle 011 \rangle$ zone axis of Form I of paracetamol, confirming that the crystal was Form I of paracetamol. No acceptable fits using Forms II and III were possible. The bulk sample from which the crystal was taken was confirmed to be Form I of paracetamol by PXRD.

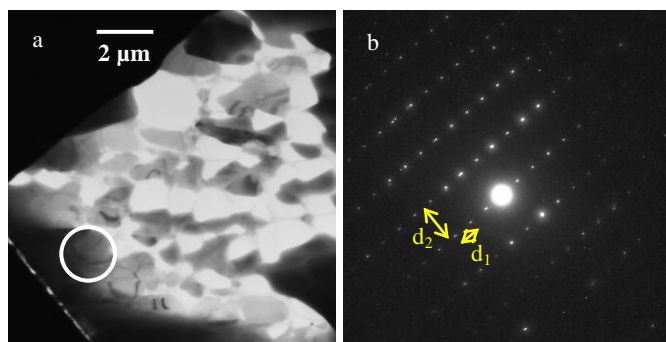


Figure 1. Transmission electron microscopy analysis of a sample of paracetamol grown from the melt. (a) TEM image showing crystallites of paracetamol which grew after molten paracetamol was spread over a TEM sample grid and allowed to cool. The circled area is the region chosen for electron diffraction analysis by use of a selected area aperture. (b) Electron diffraction pattern in agreement with the $\langle 011 \rangle$ zone axis of Form I of paracetamol. The distances d_1 and d_2 , when averaged across the pattern, were used to calculate the d-spacings of two low index crystal planes that were diffracting when the pattern was recorded.

Phase identification with Form II of paracetamol

A sample of paracetamol was prepared by heating amorphous paracetamol to 125 °C to induce crystallisation (the amorphous phase having been generated by melting paracetamol at 190 °C and then quench cooling). This preparation method has previously been found to generate Form II of paracetamol.^[29] Electron diffraction patterns were recorded from a crystal in this sample (Figure 2), and were indexed (using measured d-spacings and angles of $d_1 = 11.82$ Å, $d_2 = 8.82$ Å and $\theta = 89.8$ ° for pattern 2a and $d_1 = 17.22$ Å, $d_2 = 4.61$ Å and $\theta = 89.2$ ° for pattern 2b) to the $\langle 100 \rangle$ zone axis and $\langle 210 \rangle$ zone axis of Form II of paracetamol confirming that this crystal was Form II of paracetamol. PXRD analysis was used to confirm that the bulk sample was Form II.

Phase identification with *scyllo*-inositol

TEM analysis was performed on a fragment of a crystal of *scyllo*-inositol, a compound where the two known polymorphs have been observed to crystallise concomitantly,^[22] prepared by gently crushing crystals between two glass slides. An image and electron diffraction pattern recorded from a crystal fragment are shown in Figure 3. Attempts were made to index the diffraction pattern against the structures of Forms I and II of *scyllo*-inositol (CSD structures EFURIH05 and EFURIH04 respectively)^[22]. The pattern

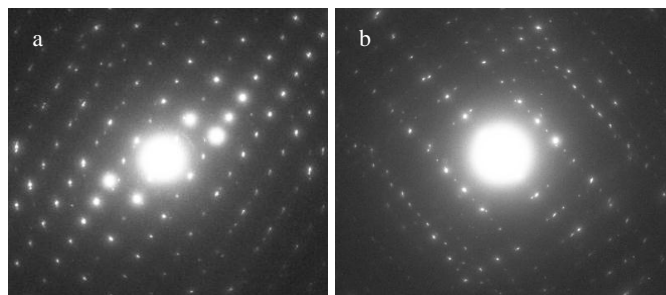


Figure 2. Electron diffraction patterns of (a) the $\langle 100 \rangle$ zone axis and (b) the $\langle 210 \rangle$ zone axis of Form II of paracetamol. In both patterns there are reflections from more than one crystal domain giving rise to a superposition of slightly differently oriented reflections.

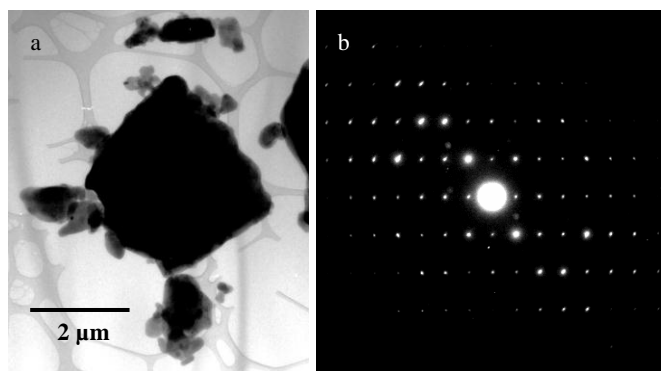


Figure 3. TEM image and electron diffraction pattern from a fragment of a crystal of *scyllo*-inositol.

was found to be consistent exclusively with diffraction from the $\langle 100 \rangle$ zone axis of Form I of *scyllo*-inositol (using measured d-spacings and angles of $d_1 = 10.70$ Å, $d_2 = 6.59$ Å and $\theta = 89.8$ °).

Phase identification with Form II of theophylline

Electron diffraction patterns (shown in Figure 4) were obtained from a crystal of theophylline prepared by cooling a saturated solution of theophylline in nitromethane. These patterns were indexed (using measured d-spacings and angles of $d_1 = 8.53$ Å, $d_2 = 3.80$ Å and $\theta = 89.9$ ° for pattern 4a and $d_1 = 8.07$ Å, $d_2 = 3.79$ Å and $\theta = 87.6$ ° for pattern 4b) to the $\langle 110 \rangle$ and $\langle 111 \rangle$ zone axes of Form II of theophylline respectively (by comparison with CSD structure BAPLOT01)^[30]. PXRD analysis was used to confirm that the bulk sample was Form II of theophylline.

Crystal structure determination using electron diffraction and crystal structure prediction

As described in the introduction, the combination of electron diffraction and crystal structure prediction could, in theory, be the basis of a new approach to crystal structure determination of unknown crystal forms.

In order to establish the viability of this approach the electron diffraction data used above for phase identification with crystals of paracetamol, *scyllo*-inositol and theophylline were compared with sets of putative crystal structures of these compounds generated by

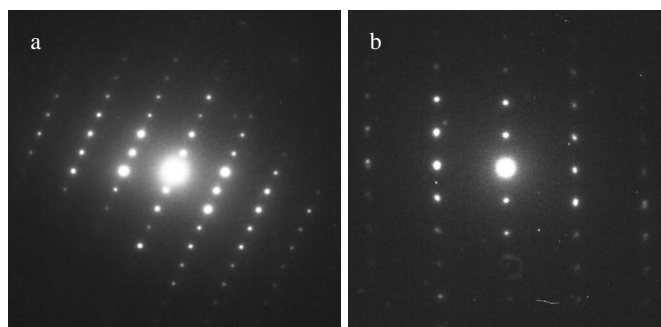


Figure 4. Electron diffraction patterns of (a) the $\langle 110 \rangle$ zone axis and (b) the $\langle 111 \rangle$ zone axis of Form II of theophylline.

Table 1. Unit cell parameters for the 14 crystal structures of paracetamol generated by CSP that have lattice energies within 10 kJ.mol⁻¹ of the global minimum. Reduced cell parameters for experimental structures of Forms I and II of paracetamol are also given (values taken from reference 31).

CSP Structure	<i>a</i> / Å	<i>b</i> / Å	<i>c</i> / Å	α / °	β / °	γ / °	Space group
AM30	12.119	8.944	7.278	90	80.03	90	<i>P2₁/n</i>
CB47	7.382	12.086	17.249	90	90	90	<i>Pbca</i>
AI22	8.248	6.749	16.669	90	126.24	90	<i>P2₁/c</i>
AY8	15.863	4.522	10.692	90	90	90	<i>Pca2₁</i>
CC19	31.121	5.074	9.886	90	82.23	90	<i>P2₁</i>
AQ6	7.344	16.281	6.571	90	90	90	<i>P2₁2₁2₁</i>
AK6	16.050	5.072	9.648	90	79.14	90	<i>P2₁/c</i>
AQ14	7.071	15.834	6.776	90	90	90	<i>P2₁2₁2₁</i>
CB9	7.266	12.207	17.432	90	90	90	<i>Pbca</i>
CC8	6.848	16.664	13.491	90	90	90	<i>Pbca</i>
AK22	15.834	5.071	9.861	90	86.68	90	<i>P2₁/c</i>
AI16	7.553	8.113	12.716	90	103.52	90	<i>P2₁/c</i>
AM4	5.937	7.590	17.071	90	80.71	90	<i>P2₁/c</i>
AK4	19.035	5.294	8.080	90	101.56	90	<i>P2₁/c</i>
(Form I)	7.085	9.370	11.706	90	97.5	90	<i>P2₁/a</i>
(FormII)	7.405	11.831	17.156	90	90	90	<i>Pbca</i>

CSP where the computer-generated structures corresponding to the experimentally observed crystal forms had been previously identified. Importantly, as the electron diffraction data was recorded from known crystal forms, and it was known which of the computationally generated structures corresponds to each of these forms, it was possible to test that the approach consistently identifies the correct structure. A demonstration that the combination of electron diffraction and crystal structure prediction could potentially be applicable for crystal structure determination would be that electron diffraction data from a particular crystal form could be indexed uniquely to the computationally generated structure that corresponds to that crystal form.

Crystal structure determination with Form I of paracetamol

CSP for paracetamol has been performed by Beyer *et al.*, who generated a set of 14 potential crystal structures with lattice energies within 10 kJ.mol⁻¹ of the global minimum (this is the energy range within which polymorphic forms of compounds are usually found).^[31] The structures, whose unit cell parameters are

Table 2. Zone axes of the 14 computationally generated paracetamol structures that could be matched to the diffraction pattern of Form I of paracetamol obtained from the melt shown in Figure 1b.

Pattern	CSP structure	Zone axis	D-spacings and Interplanar Angle		
			<i>d</i> ₁ / Å (plane)	<i>d</i> ₂ / Å (plane)	θ_1 / °
Figure 1b	(Measured values)		11.51	5.65	85.4
(Match 1)	AM30	<011>	11.936 (100)	5.593 (01-1)	82.2

summarised in Table 1, are referred to by the coordination group and number from the MOLPAK structure generation routine.^[32] The lowest energy predicted structure (AM30) was found to match Form I of paracetamol, and the second lowest in energy (CB47) to match Form II, both with good accuracy (the greatest deviation in predicted lattice parameters compared to room temperature structures being 4.5%, see Table 1).

The electron diffraction pattern shown in Figure 1b was then indexed against the 14 computationally derived potential crystal structures for paracetamol from Beyer *et al.*,^[31] taking into account d-spacings and the interplanar angle measured from this pattern, and systematic absences, using wide tolerances of +/- 5 % for d-spacings and +/- 4 ° for the interplanar angle. Only one of the structures had a zone axis which was consistent with the electron diffraction data. The respective d-spacings and interplanar angle of this matching zone axis are shown in Table 2 and a simulated electron diffraction pattern is shown in Figure 5. This structure was AM30, the predicted structure found by Bayer *et al* to match Form I.

On the basis of the electron diffraction data, therefore, putative structure AM30 was uniquely identified from the set of computationally generated paracetamol structures as matching the crystal form being analysed by TEM. If, hypothetically, the crystal structure of Form I of paracetamol had not been known in advance,

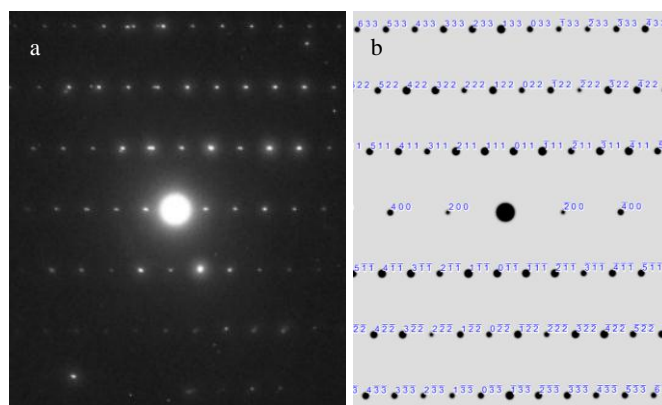


Figure 5. (a) Electron diffraction pattern of the <011> zone axis of Form I of paracetamol. (b) Simulated electron diffraction pattern from the <011> zone axis of the computationally derived paracetamol crystal structure AM30. Note that due to multiple scattering reflections are observed in positions where there are systematic absences in the simulated pattern (the simulated pattern is based on a kinematic model), and also that some differences in reflection intensities are to be expected as the computationally derived Form I structure is not identical to that observed experimentally.

this analysis would have enabled the structure to have been proposed. This result demonstrates that the combination of CSP and electron diffraction is a viable strategy for the identification of unknown crystal forms. An indication of why such an approach to crystal structure determination may be of value can be seen in Figure 1a. The electron diffraction analysis was performed on a selected section of a crystallite of paracetamol (a circular area of diameter 2 μm where the crystal thickness was approximately 0.3 μm), a region corresponding to ~ 1 pg of sample. Indeed, sub-micron sized areas, and sub-picogram amounts, of sample are routinely characterised with transmission electron microscopy. This amount of material is too small for analysis with techniques which are traditionally used for crystal structure determination and solid-state characterisation of molecular materials such as single crystal X-ray diffraction (even with a synchrotron source), powder X-ray diffraction, Raman spectroscopy and differential scanning calorimetry.

It should be noted that the tolerances which were used when attempting to match reflections in the experimental diffraction pattern to crystal planes in the computationally derived crystal structures of paracetamol are wider than would typically be used when attempting to index a diffraction pattern to an experimentally derived crystal structure. Wide tolerances were required because of the typical structural discrepancies between observed crystal structures and the nearest local minimum on the calculated lattice energy surface.^[33, 34] These errors result in part from limitations in computational models used to represent inter- and intra-molecular interactions in the crystal; all predicted crystal structures considered in the present study were energy minimised using interatomic potentials. Also contributing to the slight differences between predicted and observed structures is the comparison between a temperature-less lattice energy minimum and a measured structure at real temperature. Root-mean-squared deviations in atomic positions between predicted and observed crystal structures, as measured for a reference molecule and its 14 or 15 nearest neighbours, are typically on the order of 0.1-0.3 \AA , so that the predicted and true structures will have the same overall arrangement of molecules, intermolecular interactions, and space group.^[35, 36]

Indeed, the extent to which corresponding computationally and experimentally derived crystal structures differ is dependent on factors such as the model used during crystal structure prediction and the nature of the crystal form being predicted. For example, the paracetamol calculations were performed with completely rigid-molecule predictions, an assumption which limits the accuracy of the resulting structures. In contrast, with the *scyllo*-inositol and theophylline examples described below, flexibility of the molecule was accounted for during the energy minimisations (although differently for those two molecules), and it was therefore possible to use narrower tolerances when attempting to index electron diffraction patterns of these compounds to CSP structures. In addition, the match between corresponding known X-ray and calculated CSP structures of Form I of paracetamol is not as good as that between Form II structures. The effect of these differences can be seen in the simulated PXRD traces of the computationally and experimentally derived structures of Form I and Form II of paracetamol (Figure 6) where the traces of Form II are more closely matched than those of Form I. The magnitude of these differences will govern the ease with which electron diffraction patterns can be matched to the correct computationally derived structure when using a combined TEM/CSP approach for crystal structure determination.

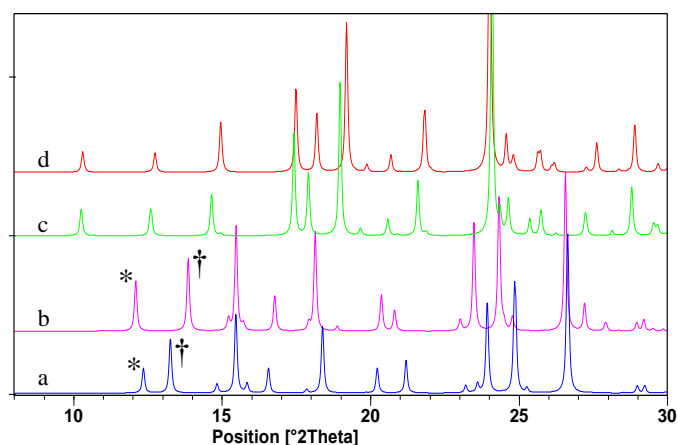


Figure 6. Comparison of simulated PXRD traces of computationally and experimentally derived crystal structures of Form I and Form II of paracetamol (the simulations assume Cu K α radiation with a wavelength of 1.5406 \AA). (a) Computationally derived Form I, AM30.^[31] (b) Experimentally derived Form I, HXACAN01.^[26] The peaks in these traces that are marked with an asterisk are due to diffraction from the same crystal planes, as are the peaks marked with a \dagger symbol. (c) Computationally derived Form II, CB47.^[31] (d) Experimentally derived Form II, HXACAN23.^[27]

Crystal structure determination with Form II of paracetamol

The two electron diffraction patterns of Form II of paracetamol shown in Figure 2 were indexed against the 14 computationally derived crystal structures for paracetamol using tolerances of $\pm 4\%$ for d-spacings and $\pm 3^\circ$ for the interplanar angle. Three of the computationally derived structures had a zone axis which was consistent with the electron diffraction pattern in Figure 2a, and two of the computationally derived structures had a zone axis which was consistent with the pattern in Figure 2b (sets of symmetry equivalent zone axes were counted as one match). The respective d-spacings and interplanar angles of these possible matches are shown in Table 3.

It can be seen from Table 3 that two of the 14 computationally generated crystal structures of paracetamol, CB47 and CB9, are consistent with both diffraction patterns of Form II shown in Figure 2. On the basis of the electron diffraction patterns alone it is not possible to determine unambiguously which of these two crystal forms corresponds to the sample being analysed. It is, however, possible to use additional information generated during the crystal structure prediction process, such as lattice energy and density, to identify which of the two forms is more likely to be observed experimentally. CB47 has a significantly lower calculated lattice energy than CB9 ($-106.5 \text{ kJ}\cdot\text{mol}^{-1}$ compared to $-101.8 \text{ kJ}\cdot\text{mol}^{-1}$), making it the more likely to be observed, and this is indeed the computationally derived structure corresponding to Form II of paracetamol.

Interestingly, there is a strong similarity between Form II of paracetamol (and also, therefore, CB47) and CB9. The structures are polytypes, sharing the same planar layer type arrangement of paracetamol molecules formed by a 2D hydrogen bond network, and only differ in that adjacent layers in CB9 are translated by half a unit cell along the b-axis in comparison with adjacent layers in Form II (Figure 7). Indeed, it is possible that Form II crystals could contain stacking faults where locally the structure is that of CB9. It may be that in future work such domain-like character might be observed within the corresponding TEM images.

Table 3. Zone axes of the 14 computationally generated paracetamol structures that could be matched to the diffraction patterns of Form II of paracetamol prepared by crystallisation from the amorphous phase shown in Figure 2.

Pattern	CSP structure	Zone axis	D-spacings and Interplanar Angle		
			$d_1 / \text{\AA}$ (plane)	$d_2 / \text{\AA}$ (plane)	$\theta_1 / ^\circ$
Figure 2a	(Measured values)		11.82	8.82	89.8
1.	AM30	<001>	11.936 (100)	8.944 (010)	90.0
2.	CB47	<100>	12.086 (010)	8.624 (002)	90.0
3.	CB9	<100>	12.162 (010)	8.725 (002)	90.0
Figure 2b	(Measured values)		17.44	4.61	89.2
1.	CB47	<210>	17.249 (001)	4.676 (1-20)	90.0
2.	CB9	<210>	17.451 (001)	4.673 (1-20)	90.0

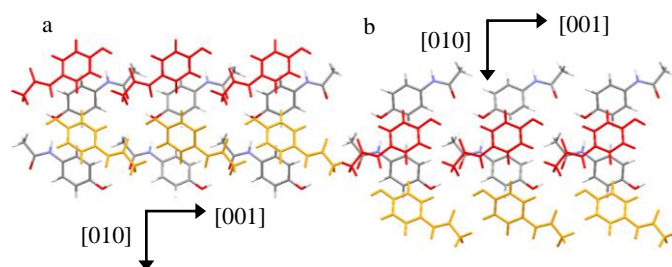


Figure 7. The computationally derived crystal structures (a) CB47 (Form II of paracetamol) and (b) CB9 viewed such that the layers formed by 2D hydrogen bonded networks of paracetamol molecules are in the plane of the page. The arrangement of adjacent layers in the two structures differs by a translation of half a unit cell along the b-axis.

Crystal structure determination with Form I of *Scyllo*-Inositol

A set of computationally derived *scyllo*-inositol crystal structures has been reported by Day *et al.*,^[22] and the unit cell parameters of the nine crystal structures with lattice energies within 15 kJ.mol⁻¹ of the global minimum are listed in Table 4. S1 corresponds to the experimentally determined Form I of *scyllo*-inositol, CSD ref. EFURIH05,^[22] while S9 corresponds to the reported Form II, EFURIH04.^[22]

Attempts were made to index the diffraction pattern shown in Figure 3b against the nine computationally derived *scyllo*-inositol crystal structures using tolerances of $d \pm 4\%$ and $\theta \pm 3^\circ$. The pattern was unambiguously matched to structure S1, that corresponding to Form I of *scyllo*-inositol (Table 5). No other predicted structures had zone axes that are consistent with the observed diffraction pattern.

Table 4. Unit cell parameters for the nine lowest energy computationally derived crystal structures of *scyllo*-inositol.

CSP Structure	$a / \text{\AA}$	$b / \text{\AA}$	$c / \text{\AA}$	$\alpha / ^\circ$	$\beta / ^\circ$	$\gamma / ^\circ$	Space group
S1	5.001	6.437	11.420	90	115.43	90	$P2_1/c$
S2	12.012	12.012	4.270	90	90	120	$R-3$
S3	21.484	6.291	11.480	90	71.10	90	$C2/c$
S4	20.173	6.352	11.427	90	97.43	90	$C2/c$
S5	9.565	6.289	11.486	90	82.22	90	$P2_1/c$
S6	8.768	7.941	9.564	90	147.23	90	$P2_1/c$
S7	8.770	7.939	9.566	90	147.24	90	$P2_1/c$
S8	11.011	6.876	11.127	90	124.78	90	$P2_1/c$
S9	6.446	6.856	8.145	92.76	104.84	94.46	$P-1$

Table 5. Zone axes of the nine computationally derived *scyllo*-inositol structures that could be matched to the diffraction pattern of *scyllo*-inositol shown in **Error! Reference source not found.**3b.

Pattern	CSP structure	Zone axis	D-spacings and Interplanar Angle		
			$d_1 / \text{\AA}$ (plane)	$d_2 / \text{\AA}$ (plane)	$\theta_1 / ^\circ$
Figure 3b	(Measured values)		10.70	6.59	89.8
(Match 1)	S1	<100>	10.314 (001)	6.437 (010)	90.0

Crystal structure determination with Form II of theophylline

A set of computationally derived theophylline crystal structures was generated by CSP, and the second lowest energy structure, #2, was found to be a match to the experimental crystal structure of the thermodynamically stable polymorph of theophylline, Form II. The 12 lowest energy predicted structures were used in this study, the unit cell parameters of which are listed in Table 6.

Attempts were made to index the two theophylline Form II diffraction patterns shown in Figure 4 against the 12 computationally derived theophylline crystal structures using tolerances of $d \pm 4\%$ and $\theta \pm 3^\circ$. The diffraction patterns were consistent with two of the structures, #2 and #4. Structure #2, the computationally derived structure corresponding to Form II of theophylline, is slightly lower in energy than #4, has a greater density and gives reflection intensities that more closely agree with observed ones, and so is the structure more likely to be a match for the sample under analysis. However, it would be difficult to

completely discount structure #4 as a match for the observed data. In this example, electron diffraction did again identify the

Table 6. Unit cell parameters for the 12 lowest energy computationally derived crystal structures of theophylline and the experimentally derived crystal structure of Form II of theophylline (BAPLOT01)^[30].

CSP Structure	<i>a</i> / Å	<i>b</i> / Å	<i>c</i> / Å	α / °	β / °	γ / °	Space group
#1	7.939	9.679	10.768	90	94.28	90	<i>P2₁/c</i>
#2	24.659	3.902	8.704	90	90	90	<i>Pna2₁</i>
#3	12.976	9.756	6.908	90	99.35	90	<i>Pbca</i>
#4	3.895	26.278	9.383	90	113.28	90	<i>P2₁/c</i>
#5	3.901	25.983	8.390	90	95.70	90	<i>P2₁/c</i>
#6	8.672	13.129	7.146	90	93.95	90	<i>Cc</i>
#7	13.076	7.003	17.783	90	90	90	<i>P2₁/c</i>
#8	9.749	6.908	25.688	90	90	90	<i>P-1</i>
#9	23.155	3.957	9.421	90	90	90	<i>Pbca</i>
#10	7.425	7.572	8.472	65.31	67.82	77.98	<i>C2/c</i>
#11	19.518	9.715	10.551	90	119.48	90	<i>C2/c</i>
#12	3.941	27.052	8.795	90	109.88	90	<i>Pna2₁</i>
(Form II)	24.612	3.830	8.501	90	90	90	<i>Pna2₁</i>

computationally derived structure corresponding to the sample being analysed, but not unambiguously.

Structures #2 and #4 do in fact show a significant difference in one unit cell dimension (the a-axis of #2 / b-axis of #4), and an electron diffraction pattern which showed diffraction from the corresponding crystal planes would readily distinguish between the two forms. Unfortunately, crystals of theophylline have a pronounced plate-like morphology, with this particular axis perpendicular to the plate face, and are always oriented on a TEM sample grid (when flat) with this axis pointing close to parallel to the electron beam. Because Bragg scattering angles are small for diffraction of electrons, and the tilt range in a TEM instrument is limited to ~ 60 °, it is not possible to observe diffraction from these planes under our experimental conditions. This limitation associated with crystal habit could be overcome by employing a different sample preparation method which avoided the preferred orientation of crystals on the sample grid.

Discussion

There are some further limitations to a combined CSP/TEM approach for crystal structure determination. Firstly, samples are held under vacuum during TEM analysis, and certain materials such

as hydrated and solvated crystal forms are likely to be unstable under these conditions. Additionally, the approach is reliant on the computational methods generating a putative crystal structure which is a good enough representation of the observed crystal structure. The CSP methodology must, therefore, generate as complete a set of crystal structures as possible, and the model used for energy minimisation should result in realistic geometries. Such structure prediction calculations are now routinely successful in generating matches for experimentally observed structures of small, rigid molecules,^[35] at least where the molecule does not crystallize in an unusual space group, or with several independent molecules in the asymmetric unit (high *Z'*). However, reliable crystal structure prediction for molecules with conformational flexibility remains a challenge,^[35] due both to the more complex structural space that must be searched and to difficulties in simultaneously modelling differences in inter- and intra-molecular contributions to the lattice energy with the required accuracy. Such challenges are being overcome^[37] and the power of combining such modelling methods with TEM will increase as CSP methodologies continue to develop.

Conclusion

It has been shown, using the pharmaceutical compounds paracetamol, *scyllo*-inositol and theophylline, that electron diffraction is a useful tool for performing polymorphic phase identification on individual crystallites in a given sample of material. Furthermore, it was demonstrated that if the crystal structures of the analysed materials had not been known prior to the analysis, the correct crystal structures could have been proposed by combining electron diffraction data with the results of crystal structure prediction. A TEM/CSP approach to crystal structure determination could find applications in situations where traditional X-ray based approaches to crystal structure determination are not applicable, for example, where only a trace amount of material is available for analysis, or where the crystal phase of interest is present as a mixture with other forms.

Experimental Section

All chemicals were purchased from Sigma-Aldrich and used as received.

Transmission electron microscopy characterization was performed at room temperature on a Philips CM30 instrument operating at 300 kV and data were collected on photographic films which were scanned in order to generate digital images. Samples were supported on holey-carbon films on 300 mesh copper grids held within a double tilt sample holder. The resulting electron diffraction patterns were indexed by comparison with predicted low energy crystal structures: the positions of reflections in experimental diffraction patterns were measured, converted to d-spacings and, along with interplanar angles, matched to calculated values for these predicted structures. The experimental diffraction patterns were then compared with simulated diffraction patterns of the appropriate zone axes to ensure a match. The simulations were carried out using CrystalMaker SingleCrystal v2.2.3 software which performs a rapid, kinematic calculation of the diffraction pattern.

Powder X-ray diffraction analysis was performed on a Philips X'Pert Diffractometer equipped with an X'celerator RTMS detector using Cu K α radiation at a wavelength of 1.5406 Å. Data were collected at ambient temperature between 3 and 50 °2 θ using a step size of 0.0167 °2 θ and a collection time of 5 minutes.

Crystal structure prediction: For paracetamol and *scyllo*-inositol, predicted structures were taken from previously published studies (Beyer *et al* for paracetamol and Day *et al* for *scyllo*-inositol).^[22, 31] Putative crystal structures of theophylline were generated with the CrystalPredictor program,^[38] using a rigid molecular geometry derived from a quantum mechanical calculation on the isolated molecule, followed by quasi-random sampling of unit cell dimensions, molecular positions and orientations within 25 common space groups (*P1*, *P-1*, *P2₁*, *P2₁/c*, *P2₁2₁2*, *P2₁2₁2₁*, *Pna2₁*, *Pca2*, *Pbca*, *Pbcn*,

C2/c, Cc, C2, Pc, P2/c, C222₁, Fdd2, Pccn, P4₁, I4₁/a, P4₁2₁2, P3₁, R-3, P3₁2₁ and P6₁), all with Z=1. The resulting crystal structures were then re-optimised using the program DMACRYS^[39] with intermolecular interactions described by an empirically parameterised exp-6 repulsion-dispersion potential (the FIT potential described by Coombes *et al.*^[33]) and an atomic multipole electrostatic model, with multipoles derived from a distributed multipole analysis^[40] of the calculated molecular charge density. Calculated lattice energies were found to be particularly sensitive to the orientation of the two methyl groups, so initial crystal structures were generated with two different orientations of the methyl hydrogen atoms; these two sets of structures were then merged and all structures within 10 kJ mol⁻¹ of the lowest energy structure were further optimised, using the CrystalOptimizer method,^[41] which combines a quantum mechanical treatment of the intramolecular energy with the atom-atom model of intermolecular interactions. Using this approach, the torsion angles determining the orientation of the two methyl groups were optimised in all crystal structures. The optimised crystal structures were then clustered to remove duplicates. Molecular energies and charge densities throughout the predictions were calculated at the B3LYP/6-31G(d,p) level of theory.

Acknowledgements

The authors thank the ERC, the Royal Society, the Pfizer Institute for Pharmaceutical Materials Science, the EPSRC and the EU INTERREG IVA 2 Mers-Seas-Zeeën Cross-border Cooperation Programme for funding. EGB also thanks the University of Cambridge Institute of Continuing Education for their support.

- [1] a) W. P. Jensen, G. J. Palenik and I. Suh, *J. Chem. Edu.* **2003**, *80*, 753-761; b) M. G. Rossmann, *Acta Crystallogr., Sect. A Found. Crystallogr.* **1998**, *A54*, 716-728.
- [2] K. D. M. Harris and E. Y. Cheung, *Chem. Soc. Rev.* **2004**, *33*, 526-538.
- [3] S. Datta and D. J. Grant, *Nat. Rev. Drug Discov.* **2004**, *3*, 42-57.
- [4] K. D. M. Harris, M. Tremayne and B. M. Kariuki, *Angew. Chemie, Int. Ed.* **2001**, *40*, 1626-1651.
- [5] D. L. Dorset, *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1998**, *A54*, 750-757.
- [6] T. E. Weirich, R. Ramlau, A. Simon, S. Hovmoeller and X. Zou, *Nature* **1996**, *382*, 144-146.
- [7] J. M. Thomas and P. A. Midgley, *Chem. Commun.* **2004**, 1253-1267.
- [8] D. A. Muller, L. F. Kourkoutis, M. Murfitt, J. H. Song, H. Y. Hwang, J. Silcox, N. Dellby and O. L. Krivanek, *Science* **2008**, *319*, 1073-1076.
- [9] D. J. H. Cockayne, I. L. F. Ray and M. J. Whelan, *Phil. Mag.* **1969**, *20*, 1265-1270.
- [10] J. P. Morniroli, A. Redjaimia and S. Nicolopoulos, *Ultramicroscopy* **2007**, *107*, 514-522.
- [11] A. Eggeman, T. White and P. Midgley, *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2009**, *A65*, 120-127.
- [12] R. E. Dunin-Borkowski, M. R. McCartney, R. B. Frankel, D. A. Bazylinski, M. Posfai and P. R. Boseck, *Science* **1998**, *282*, 1868-1870.
- [13] P. A. Midgley and M. Weyland, *Ultramicroscopy* **2003**, *96*, 413-431.
- [14] H. Sawada, Y. Tanishiro, N. Ohashi, T. Tomita, F. Hosokawa, T. Kaneyama, Y. Kondo and K. Takayanagi, *J. Electron Microsc.* **2009**, *58*, 357-361.
- [15] W. Jones, J. M. Thomas, J. O. Williams and L. W. Hobbs, *J. Chem. Soc., Faraday Trans. 2* **1975**, *71*, 138-145.
- [16] U. Kolb, T. E. Gorelik, E. Mugnaioli and A. Stewart, *Polym. Rev.* **2010**, *50*, 385-409.
- [17] A. S. Eggeman and P. A. Midgley in *Precession Electron Diffraction, Vol. 170* (Ed. P. W. Hawkes), Elsevier Ltd., **2012**, pp. 1-63.
- [18] D. B. Williams and C. B. Carter, *Transmission electron microscopy. A textbook for materials science*, Plenum Press, New York, **1996**.
- [19] a) U. Kolb, T. Gorelik and E. Mugnaioli, *Mater. Res. Soc. Symp. Proc.* **2009**, *1184*; b) M. Feyand, E. Mugnaioli, F. Vermoortele, B. Bueken, J. M. Dieterich, T. Reimer, U. Kolb, D. de Vos and N. Stock, *Angew. Chem., Int. Ed.* **2012**, *51*, 10373-10376; c) T. Gorelik, M. U. Schmidt, J. Bruning, S. Beko and U. Kolb, *Cryst. Growth Des.* **2009**, *9*, 3898-3903.
- [20] G. M. Day, *Crystallogr. Rev.* **2011**, *17*, 3-52.
- [21] a) A. T. Hulme, S. L. Price and D. A. Tocher, *J. Am. Chem. Soc.* **2005**, *127*, 1116-1117; b) A. J. Cruz Cabeza, G. M. Day, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.* **2007**, *7*, 100-107; c) K. M. Lutker and A. J. Matzger, *J. Pharm. Sci.* **2010**, *99*, 794-803.
- [22] G. M. Day, J. Van de Streek, A. Bonnet, J. C. Burley, W. Jones and W. D. S. Motherwell, *Cryst. Growth & Des.* **2006**, *6*, 2301-2307.
- [23] A. J. Cruz-Cabeza, S. Karki, L. Fabian, T. Friscic, G. M. Day and W. Jones, *Chem. Commun.* **2010**, *46*, 2224-2226.
- [24] E. Salager, G. M. Day, R. S. Stein, C. J. Pickard, B. Elena and L. Emsley, *J. Am. Chem. Soc.* **2010**, *132*, 2564-2566.
- [25] M. D. Eddleston, K. E. Hejczyk, E. G. Bithell, G. M. Day and W. Jones, *Chem. - Eur. J.* Submitted.
- [26] M. Haisa, S. Kashino, R. Kawai and H. Maeda, *Acta Crystallogr. B* **1976**, *B32*, 1283-1285.
- [27] T. N. Drebuschak and E. V. Boldyreva, *Z. Kristallogr.* **2004**, *219*, 506-512.
- [28] M.-A. Perrin, M. A. Neumann, H. Elmaleh and L. Zaska, *Chem. Commun.* **2009**, 3181-3183.
- [29] P. Di Martino, P. Conflant, M. Drache, J. P. Huvenne and A. M. Guyot-Hermann, *J. Therm. Anal.* **1997**, *48*, 447-458.
- [30] Y. Ebisuzaki, P. D. Boyle and J. A. Smith, *Acta Crystallogr. C* **1997**, *C53*, 777-779.
- [31] T. Beyer, G. M. Day and S. L. Price, *J. Am. Chem. Soc.* **2001**, *123*, 5086-5094.
- [32] J. R. Holden, Z. Du and H. L. Ammon, *J. Comput. Chem.* **1993**, *14*, 422-437.
- [33] D. S. Coombes, S. L. Price, D. J. Willock and M. Leslie, *J. Phys. Chem.* **1996**, *100*, 7352-7360.
- [34] G. M. Day, J. Chisholm, N. Shan, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.* **2004**, *4*, 1327-1340.
- [35] G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, 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 Crystallogr. B* **2009**, *B65*, 107-125.
- [36] D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti and I. K. Zhitkov, *Acta Crystallogr. B* **2011**, *B67*, 535-551.
- [37] a) G. M. Day and T. G. Cooper, *CrystEngComm* **2010**, *12*, 2443-2453; b) P. G. Karamertzanis, G. M. Day, G. W. A. Welch, J. Kendrick, F. J. J. Leusen, M. A. Neumann and S. L. Price, *J. Chem. Phys.* **2008**, *128*, 244701-244717; c) H. C. S. Chan, J. Kendrick and F. J. J. Leusen, *Angew. Chem. Int. Ed.* **2011**, *50*, 2979-2981; d) A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman, C. C. Pantelides, S. L. Price, P. T. A. Galek, G. M. Day and A. J. Cruz-Cabeza, *Int. J. Pharm.* **2011**, *418*, 168-178.
- [38] P. G. Karamertzanis and C. C. Pantelides, *J. Comput. Chem.* **2004**, *26*, 304-324.
- [39] S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis and G. M. Day, *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478-8490.
- [40] A. J. Stone and M. Alderton, *Mol. Phys.* **1985**, *56*, 1047-1064.
- [41] A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman and C. C. Pantelides, *J. Chem. Theory Comput.* **2011**, *7*, 1998-2016.

Received: ((will be filled in by the editorial staff))

Revised: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

