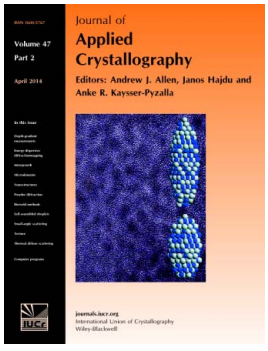
Journal of
Applied Crystallography
ISSN 1600-5767

Simultaneous determination of several crystal structures from powder mixtures: the combination of powder X-ray diffraction, band-target entropy minimization and Rietveld methods

Martin Schreyer, Liangfeng Guo, Satyanarayana Thirunahari, Feng Gao and Marc Garland

J. Appl. Cryst. (2014). **47**, 659–667



Journal of
Applied Crystallography
ISSN 1600-5767

Received 16 September 2013
Accepted 14 February 2014

Simultaneous determination of several crystal structures from powder mixtures: the combination of powder X-ray diffraction, band-target entropy minimization and Rietveld methods

Martin Schreyer,[a] Liangfeng Guo,[a] Satyanarayana Thirunahari,[b] Feng Gao[c] and Marc Garland[a]*

[a]Process Science and Modelling, Institute of Chemical and Engineering Sciences (ICES), 1 Pesek Road, Singapore 627833, [b]Crystallization and Particle Science, Institute of Chemical and Engineering Sciences (ICES), 1 Pesek Road, Singapore 627833, and [c]Heterogeneous Catalysis, Institute of Chemical and Engineering Sciences (ICES), 1 Pesek Road, Singapore 627833. Correspondence e-mail: marc_garland@ices.a-star.edu.sg

Crystal structure determination is the key to a detailed understanding of crystalline materials and their properties. This requires either single crystals or high-quality single-phase powder X-ray diffraction data. The present contribution demonstrates a novel method to reconstruct single-phase powder diffraction data from diffraction patterns of mixtures of several components and subsequently to determine the individual crystal structures. The new method does not require recourse to any database of known materials but relies purely on numerical separation of the mixture data into individual component diffractograms. The resulting diffractograms can subsequently be treated like single-phase powder diffraction data, *i.e.* indexing, structure solution and Rietveld refinement. This development opens up a host of new opportunities in materials science and related areas. For example, crystal structures can now be determined at much earlier stages when only impure samples or polymorphic mixtures are available.

1. Introduction

The first step to understanding a solid material is to know its atomic structure. In the case of materials with long-range order this means to determine the crystal structure. Generally, there are two cases to distinguish. In the first, single crystals of high quality are obtained by pure chance or experimental skill. In this case the single crystal is measured on a single-crystal diffractometer and structure solution is achieved by one of the many available methods such as direct methods (Karle & Hauptman, 1950), Patterson search (Patterson, 1935) or, more recently, charge flipping (Oszlányi & Sütő, 2004). This is followed by structure refinement using software like *SHELXL* (Sheldrick, 2008) or *JANA2006* (Petříček *et al.*, 2006). In the second case, single crystals of sufficient quality and size are not available. In this case it is necessary to optimize the synthesis and/or purification methods until single-phase powders are obtained. Obviously, this requires a lot of skill on the part of the experimentalist and typically a significant amount of time. These single-phase powders are then measured on a powder diffractometer. The data collection is followed by indexing (Visser, 1969; Werner *et al.*, 1985; Boultif & Louër, 1991; Coelho, 2003; Louër & Boultif, 2007), full pattern decomposition using the Le Bail (Le Bail, 2005) or the Pawley method (Pawley, 1981), and structure solution using slightly adapted variants of the methods used in single-crystal diffraction or simulated annealing (Kirkpatrick *et al.*, 1983). The final step is the Rietveld refinement (Rietveld, 1969) of the structure.

Often neither single crystals nor single-phase powders are available. Indeed, in many real-world situations mixtures are nearly ubiquitous. Trying to solve the crystal structures from mixtures of unknown components is generally considered an intractable problem. Nevertheless, this is the situation researchers in widely different fields like materials sciences, pharmaceutical synthesis or mineralogy frequently encounter. A few methods to distinguish phases have been suggested for very special cases of mixture analysis where line broadening (Dinnebier *et al.*, 1997) or preferred orientation (Baerlocher *et al.*, 2004) occur, but a general method that can provide individual structures from mixtures would be a major breakthrough in these areas.

The band-target entropy minimization (BTEM) algorithm (Chew *et al.*, 2002) originally analysed sets of one-dimensional spectroscopic data from mixtures in order to provide accurate estimates of the underlying pure component spectra present without recourse to any *a priori* information. This requires that the compositions vary from mixture to mixture. Naturally, this makes BTEM particularly suitable for *in situ* studies.

Table 1
Mixing table for nine mixtures for tolbutamide forms I^L, II and III in weight %.

Mixture	Form I^L (wt%)	Form II (wt%)	Form III (wt%)
1	16.7	49.8	33.5
2	45.0	20.1	34.9
3	58.4	8.4	33.2
4	16.7	16.8	66.5
5	10.0	60.2	29.9
6	44.3	44.6	11.1
7	50.2	37.5	12.4
8	33.4	33.5	33.2
9	27.1	32.2	40.7

BTEM has been widely used to study nonreactive as well as reactive systems using Fourier transform infrared (FTIR) (Widjaja *et al.*, 2002; Li *et al.*, 2003), Raman (Ong *et al.*, 2003; Sin *et al.*, 2003) and other spectroscopies. When BTEM spectral estimates are used in conjunction with spectral predictions from density functional theory, it has been possible to identify new non-isolatable species in chemical syntheses. The BTEM algorithm searches for the simplest underlying signals in a set of data. In the context of powder X-ray diffraction (PXRD) data, the algorithm searches for the simplest (lowest signal entropy) solutions associated with equation (1),

$$\hat{y} = TV^{\mathrm{T}}, \tag{1}$$

where $\hat{y}$ is a diffractogram estimate, T is a transformation vector and V^{T} is a set of right singular vectors (obtained from singular value decomposition of the mixture data). After optimization, the smoothest estimate of an individual spectrum $\hat{y}$ is obtained. For further details of the procedure see §3.

More recently, BTEM analysis has been extended to other types of data. The initial application to mixture powder X-ray diffraction data (Guo *et al.*, 2004) showed that quite accurate individual diffractograms can be obtained from five-phase systems but that a few artefacts remained in the individual diffractograms. These artefacts complicated the subsequent determination of the individual crystal structures owing to resulting problems with peak fitting and integration.

Recently, the method has been further improved and extended so that a complete data analysis of the resulting single-phase diffractograms is possible. In the following, individual crystal structures are obtained from (*a*) a three-component inorganic system, (*b*) an organic trimorphic system and (*c*) a generic ternary organic system. This is achieved by measuring the diffractograms of a few distinct mixtures, obtaining individual pure diffractograms using the BTEM algorithm, improving the single-phase diffractograms using various signal processing tools, and then processing each pure component diffractogram in exactly the same way as conventional single-phase powder diffractograms to obtain the refined crystal structures. This three-step method will be referred to as the PXRD–BTEM–Rietveld approach in subsequent sections.

2. Experimental

2.1. Experimental design

All ternary mixtures were prepared in such a way that all phases were present in each mixture in varying ratios. It was attempted to emulate a random set of concentrations. Table 1 shows the composition of the nine mixtures of tolbutamide polymorphs. Fig. 1 displays the compositions as a triangle plot. Details for the generic organic sample set (Schreyer *et al.*, 2011) and the inorganic sample set (Madsen *et al.*, 2001) are reported elsewhere.

2.2. Sample preparation and measurement

Paracetamol (Sigma, purity >99%), α-lactose monohydrate (Sigma, purity >99%) and α-glycine (Aldrich, purity >98%) were used as received from the supplier. Tolbutamide (Sigma–Aldrich, 99% purity) was recrystallized in order to obtain the desired polymorphs. Tolbutamide I^L was obtained by recrystallization from hot acetonitrile followed by cooling. Tolbutamide II was obtained by adding distilled water to a solution of tolbutamide in acetonitrile, while form III was obtained by recrystallization from ethanol. More detailed procedures are reported elsewhere (Thirunahari *et al.*, 2010).

For the organic samples the pure compounds were ball-milled for 15 min at 20 Hz in a Retsch MM200 mixer mill (Retsch GmbH, Haan, Germany) using a 25 ml stainless steel beaker and a steel ball of 10 mm diameter in order to reduce preferred orientation. The mixture samples were then prepared by weighing in the desired amounts and manually grinding them in an Agate mortar for several minutes. For the polymorphic samples tolbutamide I^L and III were ball-milled under the same conditions as above but for 10 min only. Tolbutamide II was carefully hand ground to prevent conversion to polymorph I^L. Subsequently, mixtures were prepared by weighing the desired ratios into the grinding jars

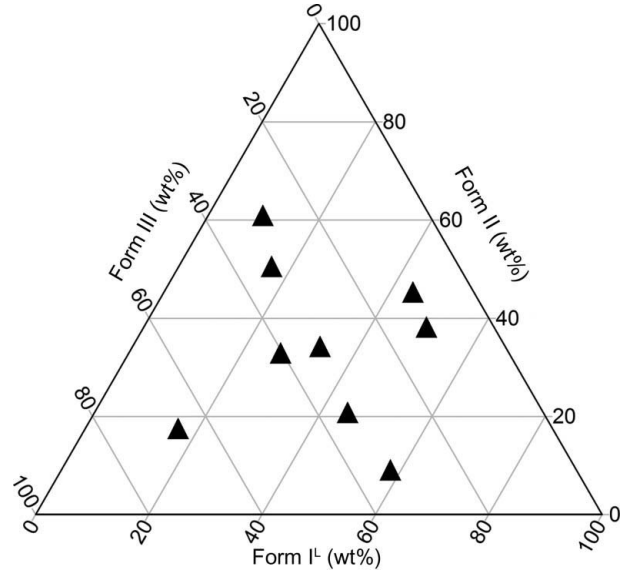


Figure 1
Triangle plot showing the compositions of the mixtures of tolbutamide forms I^L, II and III.

Table 2
Overview of experimental conditions for all PXRD experiments.

	Polymorph series	Organic series	Inorganic series†
Geometry	Bragg–Brentano	Bragg–Brentano	Bragg–Brentano
Radiation	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$
Goniometer radius	217.5 mm	217.5 mm	173 mm
Soller slits	2.5°	2.5°	4.2°
Divergence slit	0.5°	0.5°	1°
Air-scattering screen	Yes	None	None
Antiscatter slit	None	None	1°
Receiving slit	None	None	0.3 mm
Monochromator	Ni filter	Ni filter	Graphite crystal
Detector type	One-dimensional PSD (Vantec)	One-dimensional PSD (Vantec)	Zero dimensional point detector
Start 2θ	5°	12°	5°
End 2θ	50°	125°	150°
Step width	0.0167°	0.0167°	0.02°
Time per step	0.30 s	0.15 s	3 s

† The experiments for the inorganic series were not conducted by the authors but by Madsen *et al.* (2001) for the Round Robin on Quantitative Phase Analysis.

and shaking them for 10 min at 20 Hz without adding a grinding ball.

Powder X-ray diffraction data were collected in Bragg–Brentano geometry on a Bruker D8 diffractometer (Bruker AXS GmbH, Karlsruhe, Germany) equipped with a copper source (Cu $K\alpha$), a nickel filter, a nine-position sample changer and a gas-filled position-sensitive detector (PSD). All details are provided in Table 2.

3. Computational details

3.1. Single-phase diffractogram estimation using BTEM analysis

BTEM first appeared in 2002 (Chew *et al.*, 2002) and has since been used in well over 100 peer-review journal publications. As mentioned in the introduction, the goal of the BTEM algorithm is to analyse sets of spectroscopic data and thereby recover spectral estimates of the constituents without the use of any *a priori* information. The BTEM algorithm implements concepts derived from entropy minimization. It is well recognized in the literature that entropy minimization is closely associated with pattern recognition (Watanabe, 1981) and the related concept of 'the principle of simplicity' (Kapur, 1989). Accordingly, BTEM searches for the simplest underlying patterns (pure component spectra) in a spectral data set. It has been extensively used to recover one-dimensional patterns from FTIR spectroscopy (Widjaja *et al.*, 2002), Raman spectroscopy (Ong *et al.*, 2003), UV–Vis spectroscopy (Gao *et al.*, 2009), UV–Vis circular dichroism (Cheng *et al.*, 2008), FTIR microscopy (Widjaja & Tan, 2008), Raman microscopy (Widjaja & Garland, 2008), ^{1}H NMR, ^{13}C NMR (Guo *et al.*, 2008) and PXRD (Guo *et al.*, 2004), two-dimensional patterns from correlation spectroscopy, heteronuclear single quantum coherence (Guo *et al.*, 2005) and excitation fluorescence (Guo & Garland, 2007), and even three-dimensional patterns (Guo, 2006).

In the PXRD context, let $Y_{k\times\theta}$ be a set of k diffractograms where θ is the number of data channels along the 2θ axis. Then a singular value decomposition of $Y_{k\times\theta}$ provides a set of left singular vectors $U_{k\times k}$, a set of diagonal singular values $\Sigma_{k\times\theta}$ and a set of right singular vectors $V^{\mathrm{T}}_{\theta\times\theta}$ as stated in equation (2):

$$Y_{k\times\theta} = U_{k\times k}\Sigma_{k\times\theta}V^{\mathrm{T}}_{\theta\times\theta}. \quad (2)$$

The latter set $V^{\mathrm{T}}_{\theta\times\theta}$ contains the primary information concerning the intensity variation in the data set. They will be the primary mathematical objects of interest in this analysis, and as shown elsewhere (Guo *et al.*, 2004), only the first few vectors in $V^{\mathrm{T}}_{\theta\times\theta}$ will have clearly discernible localized band shape information, whereas the remaining vectors are basically noise. Accordingly, we will truncate the set of right singular vectors $V^{\mathrm{T}}_{\theta\times\theta}$ and discard the vectors containing primarily noise. The remaining set is denoted as $V^{\mathrm{T}}_{z\times\theta}$, where z represents the number of vectors retained. It is important to note that in most real systems the value of z is probably much larger than s, the number of diffracting constituents in the system.

In order to recover the underlying single-phase diffractogram estimates from an appropriate data set, equation (1) is used repeatedly to recover each estimate one at a time, until all of the information in the data set has been exhaustively searched. To implement equation (1), an entropy-related function of each updated diffractogram must be evaluated. This is normally done by evaluating the smoothness (derivatives) of the entire diffractogram point by point. The evaluated entropies of each diffractogram estimate are then used in the objective function

$$F_{\mathrm{obj}} = H + P, \quad (3)$$

where H is an entropy term and P is a penalty term (to constrain unphysical solutions, *i.e.* negative values for the estimated diffractogram). A global search is then executed using simulated annealing to find the optimum and hence each spectral estimate one at a time. For further details, the article by Widjaja *et al.* (2002) is recommended.

3.2. Crystal structure solution and refinement

Subsequently, the BTEM-separated powder patterns were treated like conventional powder patterns, with the additional difficulty that the single-phase patterns needed to be assigned to the correct substances. These BTEM-separated single-phase patterns still contained minor artefacts. Some artefacts are discernible through their unusual peak shape (sinusoidals or even negative intensities). Others can be identified as residuals of main peaks of the other single-phase patterns. Such clearly identifiable artefacts were ignored during the indexing step. Otherwise, indexing was performed like with conventional single-phase powder patterns, using the method implemented in *TOPAS 4.2* (Coelho, 2003, 2009). This was followed by full powder pattern decomposition according to the Pawley method, a so-called Pawley fit (Pawley, 1981). This step gives the extracted intensities required for structure

solution. When these fits indicate the presence of artefacts in the BTEM-estimated single-phase diffractograms, a clean-up procedure is needed. These artefacts arise from slight sample displacements of the original mixture data sets and slightly different peak asymmetries due to different packing densities. The clean-up procedure is implemented by a combined, simultaneous full powder pattern decomposition of all original mixture diffractograms. For this purpose a constrained simultaneous Pawley fit of all mixture diffractograms with all three Pawley phases was conducted. In this refinement the relative peak intensities of each phase were fixed; only variation with respect to preferred orientation along the principal axis was allowed. As an output of the combined Pawley fit, the calculated individual phase diffractograms are obtained. These reconstructed single-phase diffractograms were then used for structure solution.

Structure solution was first attempted using charge flipping (Coelho, 2007) as implemented in *TOPAS*. This method is preferred since it requires no chemical *a priori* knowledge of the composition of each phase. If charge flipping failed, direct space methods like simulated annealing as included in *TOPAS* (Coelho, 2000) or parallel tempering (Battaille *et al.*, 2006) as implemented in *Materials Studio* (Andzelm, 2007) were attempted. The main problem at this point is that it is helpful to know which phase is assigned to which molecular structure. For this purpose the method described by Scarlett & Madsen (2006) for the quantification of phases with partially or no known crystal structures (PONKCS) was used with a constraint: each phase was assigned an arbitrary unit-cell mass of 100. This gives not an absolute composition but the relative compositions of each phase in each mixture. This concentration profile can then be compared with the concentration profile obtained by a complimentary method. For example, X-ray fluorescence (XRF) can be used for inorganic samples or IR/Raman or NMR spectroscopy in the case of organic samples. Alternatively a trial and error approach can be applied, if the number of candidates is not too large. Since the direct space methods require knowledge of the molecular structure, the molecular geometry of the test molecules was first optimized in vacuum using density functional theory at the B3LYP level with *Gaussian* (Frisch *et al.*, 2004). Subsequently, with the optimized molecular geometry used as an input, the crystal structure solution was performed in *Materials Studio* using *PowderSolve* (Engel *et al.*, 1999), or with *TOPAS*. The thus obtained structures were then further refined using *TOPAS*.

4. Results and discussion

The inorganic system consisted of a series of ternary mixtures of calcium fluoride, aluminium oxide and zinc oxide. These data as well as complimentary XRF data have been used for the IUCR Round Robin on Quantitative Phase Analysis (Madsen *et al.*, 2001) and are available in the public domain (Madsen, 2001). These references also contain all experimental details. The data (Fig. 2*a*) were re-aligned, *i.e.* zero shift correction, before further processing. BTEM analysis

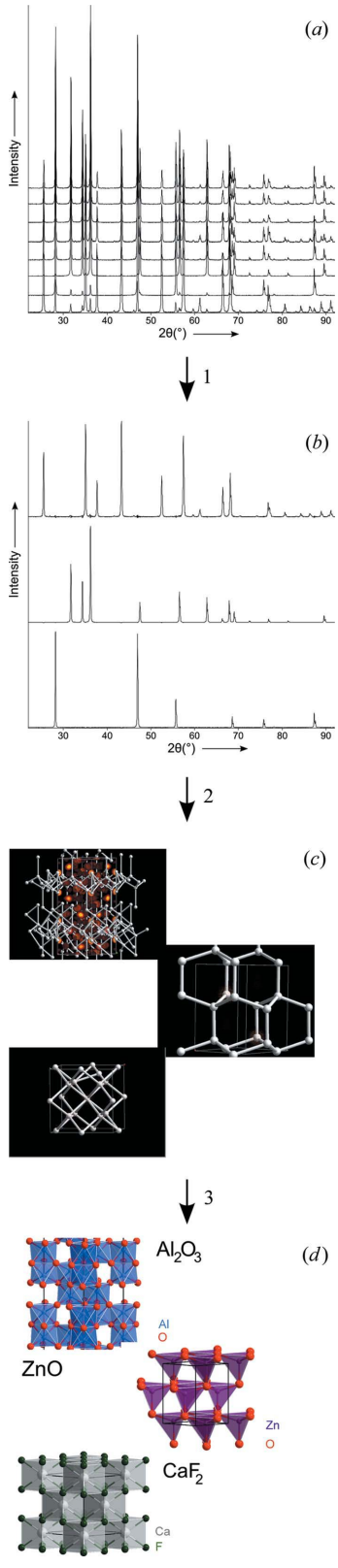


Figure 2
Schematic diagram of the primary steps in the analysis of ZnO, CaF_2 and Al_2O_3 mixtures. (*a*) The raw mixture diffractograms, (*b*) recovered 'single-phase' diffractograms, (*c*) structure models, (*d*) final crystal structures for each phase. Step 1: Application of BTEM to raw data. Step 2: Indexing, Pawley refinement and charge flipping. Step 3: Assignment of atom types and Rietveld refinement.

Table 3
Comparison of structure parameters for ZnO, CaF_2 and Al_2O_3 refined from BTEM-separated powder X-ray diffraction data and from single-phase neutron powder diffraction data.

	BTEM (Phase 1)	ICSD No. 65120
Data type	Powder XRD/ BTEM separated	Powder neutron/ single phase
Stoichiometry	ZnO	ZnO
Space group	$P6_3mc$ (186)	$P6_3mc$ (186)
a (Å)	3.24931 (2)	3.24992 (5)
c (Å)	5.20571 (4)	5.20658 (8)
M coordinate	Zn 2b $\frac{1}{3}$, $\frac{2}{3}$, $\frac{1}{2}$	Zn 2b $\frac{1}{3}$, $\frac{2}{3}$, $\frac{1}{2}$
X coordinate	O 2b $\frac{2}{3}$, $\frac{1}{3}$, 0.3790 (6)	O 2b $\frac{2}{3}$, $\frac{1}{3}$, 0.3819 (1)
$M-X$ distances (Å)	1 × 1.973 (3), 3 × 1.979 (1)	1 × 1.9884 (5), 3 × 1.9746 (2)

	BTEM (Phase 2)	ICSD No. 2908
Data type	Powder XRD/ BTEM separated	Powder neutron/ single phase
Stoichiometry	CaF_2	CaF_2
Space group	$Fm\bar{3}m$ (225)	$Fm\bar{3}m$ (225)
a (Å)	5.46329 (1)	5.462 (3)
c (Å)	= a	= a
M coordinate	Ca 4a 0, 0, 0	Ca 4a 0, 0, 0
X coordinate	F 8c $\frac{1}{4}$, $\frac{1}{4}$, $\frac{1}{4}$	F 8c $\frac{1}{4}$, $\frac{1}{4}$, $\frac{1}{4}$
$M-X$ distances (Å)	8 × 2.3656 (1)	8 × 2.365 (7)

	BTEM (Phase 3)	ICSD No. 51687
Data type	Powder XRD/ BTEM separated	Powder neutron/ single phase
Stoichiometry	Al_2O_3	Al_2O_3
Space group	$R\bar{3}c$ (167)	$R\bar{3}c$ (167)
a (Å)	4.7582 (1)	4.7597 (1)
c (Å)	12.9897 (2)	12.9935 (3)
M coordinate	Al 12c 0, 0, 0.35229 (7)	Al 12c 0, 0, 0.3523 (1)
X coordinate	O 18e 0.3062 (3), 0, $\frac{1}{4}$	O 18e 0.3065 (1), 0, $\frac{1}{4}$
$M-X$ distances (Å)	3 × 1.854 (2), 3 × 1.972 (2)	3 × 1.8535 (6), 3 × 1.9736 (9)

provided three single-phase diffractograms (Fig. 2*b*); however, the identity of the phases is not yet established. Next, each single-phase diffractogram was treated like a conventional single-phase powder diffractogram using methods traditionally employed for this purpose. All three patterns were indexed, and then unit-cell refinements were performed using the Pawley method in order to extract the intensities for structure solution. The structure solutions were achieved using charge flipping as implemented in *TOPAS* without actual knowledge of the chemistry. For each unit cell one heavy-atom position and one light-atom position was found (Fig. 2*c*), but the exact assignment of the atoms is not yet known.

At this point some quantitative analysis was required to identify the phases. In this case the relative abundance of crystalline phases obtained *via* PONKCS was correlated with the abundance of the elements found by XRF (Madsen *et al.*, 2001). This led to the identification of phase 1 as ZnO, phase 2 as CaF_2 and phase 3 as Al_2O_3. The final step was then the Rietveld refinement of each structure as shown in Fig. 2(*d*). The obtained crystallographic values of each phase are compared with values obtained from single-phase neutron powder diffraction data in Table 3. Structures from neutron powder data were chosen as a reference, because neutron diffraction is generally the most accurate powder diffraction method to establish atomic positions since light atoms such as oxygen are strong scatterers of neutrons and neutron diffraction gives stronger intensities at high angles. It can be clearly seen that the crystallographic parameters of all three phases are in good agreement with the literature results.

A frequently encountered phenomenon in the chemical process industries and specifically the pharmaceutical industry is crystal polymorphism (Bernstein, 2002). A typical example is tolbutamide, which is a sulfonyl urea drug for type II diabetic patients to control the blood sugar levels. It is known to exist in six polymorphic forms (forms I^L, I^H, II–V) (Burger, 1975; Hasegawa *et al.*, 2009; Thirunahari *et al.*, 2010; Nath & Nangia, 2011). The separation of the diffractograms of several polymorphs of a substance is one of the potential applications for the PXRD–BTEM–Rietveld approach. For this purpose, three tolbutamide polymorphs (forms I^L, II and III) were chosen. Form II is the thermodynamically most stable form, and forms I^L and III are kinetically stable forms at ambient conditions.

The measured mixture diffractograms were submitted to BTEM and three first estimates of the single-phase diffractograms were obtained as shown in Fig. 3(*b*). In this case, the BTEM estimates contained a few small artefacts. These flaws in the original data were corrected with the aid of combined, simultaneous full pattern decomposition of all original mixture diffractograms. The reconstructed single-phase diffractograms are presented in Fig. 3(*c*).

Structure solution was first attempted by charge flipping. However, it was revealed that the data quality was insufficient for this method to yield an acceptable starting model. Therefore, a direct space method, parallel tempering, was chosen as structure solution method. The crystal structure solution was performed in the *Materials Studio* environment using *PowderSolve*. Structural models for forms I^L and III were obtained within a few hours. The structure solution for form II could not be performed in an acceptable time frame because the number of degrees of freedom (DOFs) is too high. Forms I^L and III each have only a single molecule in the asymmetric unit, whereas form II contains four molecules in the asymmetric unit. As a consequence there are 50 (ten translational, 12 rotational and 28 torsional) DOFs for form II. In order to reduce the number of DOFs, a common molecular feature, *i.e.* hydrogen bonding between amine protons and carbonyl oxygen observed in tolbutamide forms I^L and III, was used to build a motif using two motion groups of tolbutamide molecules. This reduced the DOFs to 21. A further reduction was achieved by eliminating all torsional DOFs. This gave a crude structure model. During the Rietveld refinement all torsion angles were refined freely, giving the final structure solution for Rietveld refinement with *TOPAS*.

The Rietveld refinements converged to R_{wp} values of 12.2% (form I^L), 13.7% (II) and 10.4% (III). These R factors are first indicators for reasonably good structure solutions. Therefore, the resulting crystal structures were compared with structures from the Cambridge Structural Database (Allen, 2002). The refined structures agree well with the structures from the

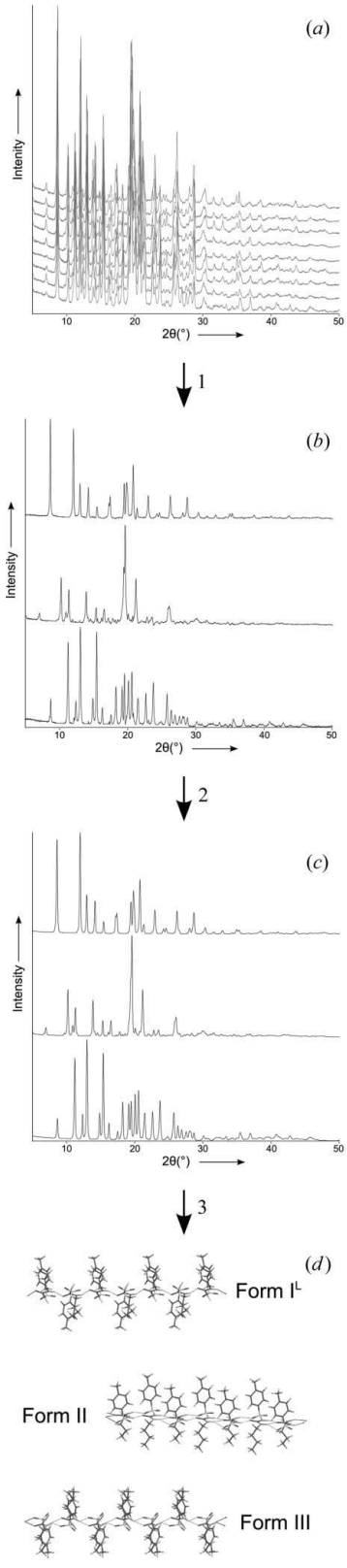


Figure 3
Schematic diagram of the primary steps in the analysis of the trimorphic tolbutamide system. (*a*) The raw mixture diffractograms, (*b*) recovered 'single-phase' diffractograms, (*c*) updated 'single-phase' diffractograms, (*d*) final crystal structures for each phase. Step 1: Application of BTEM to raw data. Step 2: Indexing, individual Pawley refinements and combined Pawley refinement of mixture data. Step 3: Parallel tempering and Rietveld refinement.

database, as shown in Fig. 4. In all three forms the same intermolecular connectivity *via* hydrogen bonding is found as in the structures from the database. In forms I^L and III tolbutamide is found in a boat conformation and in form II it is found in a chair conformation. These conformations agree with the literature. There are some minor differences in the details but this is to be expected when comparing structure solutions from powder diffraction data with structure solutions from single-crystal diffraction data.

A further organic system consisted of ten mixtures of three simple organic substances, namely paracetamol form I (Haisa *et al.*, 1976), α-lactose monohydrate (Smith *et al.*, 2005) and α-glycine (Jönsson & Kvick, 1972). The sample preparation and measurement are described elsewhere (Schreyer *et al.*, 2011). The data treatment followed a similar procedure as used for the trimorphic system. Therefore, BTEM analysis of the raw data yielded three single-phase diffractograms. These were indexed and Pawley refinements conducted. At this point, *a priori* information concerning the constituents present is introduced. The unit-cell volume of each Pawley phase is compared with the molecular volume of each candidate molecule, which can be obtained from crystallographic soft-

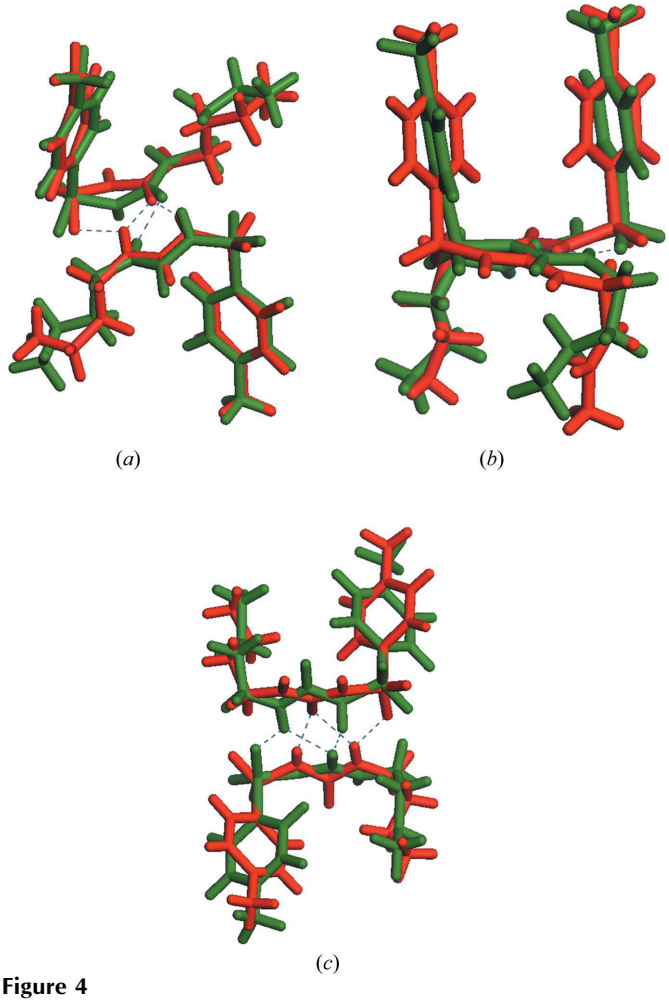


Figure 4
Superposition of structural motifs of the single-crystal structure models (red) of the three tolbutamide forms with the PXRD structure solutions (green). (*a*) Tolbutamide form I^L, (*b*) tolbutamide II, (*c*) tolbutamide III.

ware such as *Endeavour* (Putz & Brandenburg, 2012). This allows each unit cell to be assigned to the correct substance. Alternatively, a trial and error approach can be used where each candidate molecule will be tried in each unit cell. Structurally implausible solutions will have to be discarded as wrong. Accordingly, structure solution was achieved using simulated annealing as implemented in *TOPAS* followed by Rietveld refinement. Fig. 5 shows a superposition of the structure solutions from this article with literature data.

The analysis of the inorganic case proved to be quite straightforward for a number of reasons: there is little peak overlap between the different constituents, the peaks are narrow and well defined, and there is no preferred orientation for any of the constituents. Therefore, the BTEM-estimated single-phase diffractograms could be used for structure solution without further processing. In order to achieve the correct chemical assignments of the constituents, relative concentrations obtained *via* PONKCS were combined with quantitative XRF measurements of the same mixtures.

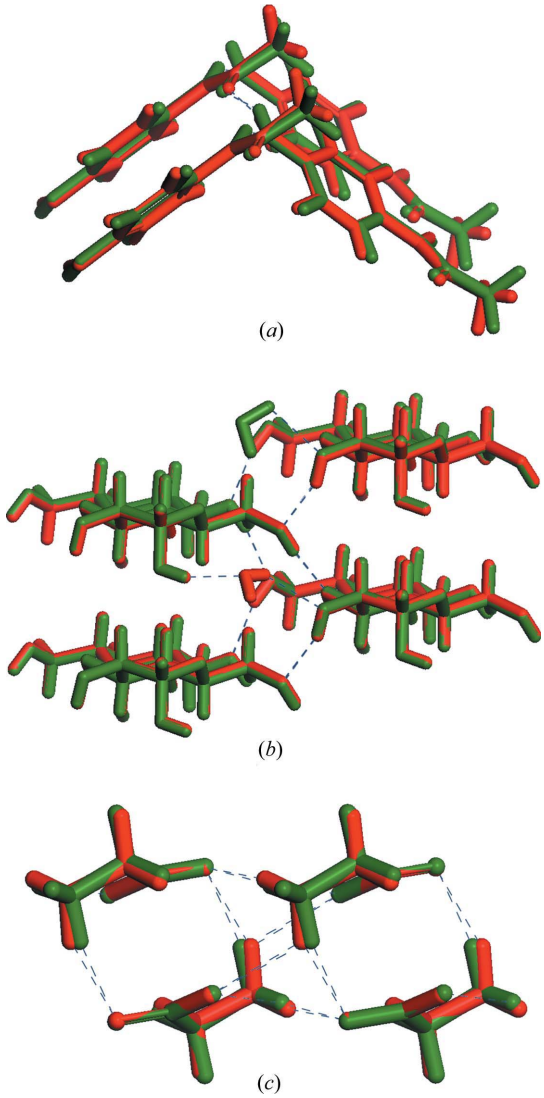


Figure 5
Superposition of structural motifs of the single-crystal structure models (red) of the three organic phases with the PXRD structure solutions (green). (*a*) Paracetamol (I), (*b*) α-lactose monohydrate, (*c*) α-glycine.

The analysis of the trimorphic organic system and the generic ternary organic system proved to be less straightforward for the following reasons. Despite careful sample preparation there was some variation, unrelated to concentration changes but due to preferred orientation. Furthermore, there was some variation in peak asymmetry due to differences in packing density. Such variations in peak intensity and shape, which are unrelated to compositional changes, are a difficulty for the BTEM method and may cause some artefacts in the BTEM estimates. This can be explained by the way the BTEM method works. The BTEM model captures the position-wise variations in intensity from mixture to mixture, and it assumes that all these position-wise intensity variations correlate linearly with compositional changes. Obviously, variation in preferred orientation directly runs counter to this assumption. The effect of variations in peak asymmetry is more subtle: in this case only the integrated area is linearly correlated with composition, while the position-wise intensity is slightly varying and thus may be misinterpreted in the BTEM model.

In order to remove the artefacts, additional steps had to be taken. A combined Pawley fit of the original raw mixture data was used to remove the artefacts in the single-phase diffractograms. Structure solution was achieved *via* parallel tempering followed by Rietveld refinement.

Finally, a comparison between the present structure solutions and Rietveld refinements from single-phase reference powder patterns of all three tolbutamide forms was performed. For this purpose the same starting models were chosen for the single-phase reference refinements. Molecular position, molecular rotation angles and torsion angles of single bonds were freely refined exactly as for the refinements of the BTEM-separated patterns. The structural similarity with single-crystal structure solution from the Cambridge Structural Database was determined by superimposing motifs from our structure refinements with equivalent motifs from single-crystal structure solutions and calculating correlation indices based on the Euclidean distance method (Basak *et al.*, 1988). The results are presented in Table 4 and show that the quality of the structure solutions obtained from the multi-component analysis is in all cases comparable to that obtained from single-phase powders. Moreover, this table confirms that the low similarity of 88.7% obtained in the multi-component analysis for tolbutamide II is not due to any shortcomings in the BTEM analysis. Indeed, single-phase analysis provides a similarity of 89.7%. The Rietveld refinements and superpositions of structures from the single-phase analyses are provided in the supplementary materials.[1]

5. Conclusion

There are currently three bottlenecks in solving crystal structures from powder diffraction data. These are (1)

[1] Supporting information for this article is available from the IUCr electronic archives (Reference: TO5064).

Table 4
Comparison of structure solutions obtained from single-phase (SP) reference powder patterns and BTEM-separated powder patterns in terms of similarity to single-crystal structures.

	SP–PXRD (%)	BTEM–PXRD (%)
Tolbutamide I^L	98.8	96.2
Tolbutamide II	89.7	88.7
Tolbutamide III	97.9	98.4

obtaining a single-phase powder, (2) indexing and (3) finding a structure model. Recently, there has been considerable progress in addressing the last two issues (Boultif & Louër, 1991; Coelho, 2003; Oszlányi & Sütő, 2004; Louër & Boultif, 2007). However, up to now no numerical approach has been proposed to address the first. In fact, this has been previously considered as purely a synthetic problem and has been left entirely to the skill of the experimentalist. The PXRD–BTEM–Rietveld method is the first purely numerical approach to address problem (1) by obtaining single-phase powder diffractograms from an ensemble of powder mixture diffractograms alone (note that BTEM analysis of such an ensemble would be tractable in spite of the small variation in line broadening from sample to sample). The current work demonstrates that the resulting and numerically recovered single-phase powder diffractograms are often suitable for *ab initio* structure solution and can be evaluated in the same way as real single-phase diffractograms. Structure solutions have been achieved not only for high-symmetry inorganic compounds but also for three different polymorphs of a low-symmetry organic compound and a ternary organic mixture. This new approach clearly offers a range of potential opportunities in mixture powder analysis of materials from a variety of sources.

Presently, the new technique has a number of limitations, most of which are related to experimental design: (1) preferred orientation needs to be negligible or at least to vary negligibly from mixture to mixture, (2) a larger number of mixture diffractograms than the number of components is required, and (3) the relative concentrations of all components need to vary so that data colinearities do not exist. Despite these experimental limitations, the BTEM-approximated single-phase patterns yielded reasonably good structure solutions for all cases discussed. If the structure solutions compare unfavourably with single-crystal structure solutions, then it is mainly because of the limitation of powder diffraction techniques, especially Bragg–Brentano diffractometry, rather than the PXRD–BTEM–Rietveld approach.

It is obvious from point (1) that Bragg–Brentano geometry is not the method of choice for such experiments; Debye–Scherrer geometry with a monochromatic source is expected to give much better results. With respect to issues (2) and (3), a particularly important challenge for future work is to obtain single-phase powder diffractograms from a single mixture, by making use of the inhomogeneity of any powder mixture on the micrometric scale. Since any multiphase powder has large inhomogeneities on the microscopic scale even when it is homogeneous on the macroscopic scale, this would be of particular interest when (1) only a few or even only one mixture is available or (2) the phases of interest are only minor components of the powders under investigation. There are severe obstacles to such an approach in microdiffraction, namely the requirement of (*a*) a bright and strongly focused X-ray source so that many independent measurements can be performed on one small sample, (*b*) good particle statistics and (*c*) high resolution in the resulting diffractograms in order to achieve structure solution. Future work will be directed to resolving the aforementioned issues.

This work was supported by the Agency for Science, Technology and Research (A*STAR), Singapore. The authors would like to thank Dr Thomas Baikie (NTU, Singapore) for insightful discussions.

References

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Andzelm, J. (2007). *Chem. World*, **4**, 72.
Baerlocher, C., McCusker, L. B., Prokic, S. & Wessels, T. (2004). *Z. Kristallogr.* **219**, 803–812.
Basak, S., Magnuson, V., Niemi, G. & Regal, R. (1988). *Discret. Appl. Math.* **19**, 17–44.
Battaille, T., Mahe, N., Le Fur, E., Pivan, J. Y. & Louër, D. (2006). *Z. Kristallogr. Suppl.* **23**, 9–14.
Bernstein, J. (2002). *Polymorphism in Molecular Crystals*. New York: Oxford University Press.
Boultif, A. & Louër, D. (1991). *J. Appl. Cryst.* **24**, 987–993.
Burger, A. (1975). *Sci. Pharm.* **43**, 161–168.
Cheng, S., Gao, F., Krummel, K. I. & Garland, M. (2008). *Talanta*, **74**, 1132–1140.
Chew, W., Widjaja, E. & Garland, M. (2002). *Organometallics*, **21**, 1982–1990.
Coelho, A. A. (2000). *J. Appl. Cryst.* **33**, 899–908.
Coelho, A. A. (2003). *J. Appl. Cryst.* **36**, 86–95.
Coelho, A. A. (2007). *Acta Cryst.* A**63**, 400–406.
Coelho, A. A. (2009). *TOPAS*. Bruker AXS GmbH, Karlsruhe, Germany.
Dinnebier, R. E., Olbrich, F., van Smaalen, S. & Stephens, P. W. (1997). *Acta Cryst.* B**53**, 153–158.
Engel, G. E., Wilke, S., König, O., Harris, K. D. M. & Leusen, F. J. J. (1999). *J. Appl. Cryst.* **32**, 1169–1179.
Frisch, M. J. *et al.* (2004). *GAUSSIAN03*. Gaussian Inc., Pittsburgh, Pennsylvania, USA.
Gao, F., Zhang, H., Guo, L. & Garland, M. (2009). *Chemom. Intell. Lab. Syst.* **95**, 94–100.
Guo, L. (2006). *Development, 2D-and 3D-BTEM for Pattern Recognition in Higher-Order Spectroscopic and Other Data Arrays.* National University of Singapore.
Guo, L. & Garland, M. (2007). *Appl. Spectrosc.* **61**, 148–156.
Guo, L., Kooli, F. & Garland, M. (2004). *Anal. Chim. Acta*, **517**, 229–236.
Guo, L., Sprenger, P. & Garland, M. (2008). *Anal. Chim. Acta*, **608**, 48–55.
Guo, L., Wiesmath, A., Sprenger, P. & Garland, M. (2005). *Anal. Chem.* **77**, 1655–1662.
Haisa, M., Kashino, S., Kawai, R. & Maeda, H. (1976). *Acta Cryst.* B**32**, 1283–1285.
Hasegawa, G., Komasaka, T., Bando, R., Yoshihashi, Y., Yonemochi, E., Fujii, K., Uekusa, H. & Terada, K. (2009). *Int. J. Pharm.* **369**, 12–18.
Jönsson, P.-G. & Kvick, Å. (1972). *Acta Cryst.* B**28**, 1827–1833.

Kapur, J. N. (1989). *Maximum Entropy Models in Science and Engineering*. New Delhi: John Wiley and Sons.
Karle, J. & Hauptman, H. (1950). *Acta Cryst.* **3**, 181–187.
Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983). *Science*, **220**, 671–680.
Le Bail, A. (2005). *Powder Diffr.* **20**, 316–326.
Li, C., Widjaja, E. & Garland, M. (2003). *J. Catal.* **213**, 126–134.
Louër, D. & Boultif, A. (2007). *Z. Kristallogr. Suppl.* **26**, 191–196.
Madsen, I. C. (2001). *IUCr CPD Round Robin on Quantitative Phase Analysis*, *Standard Powder X-ray Diffraction Data Sets – Data Analysis Kit*, http://www.mx.iucr.org/iucr-top/comm/cpd/QARR/data-kit.htm.
Madsen, I. C., Scarlett, N. V. Y., Cranswick, L. M. D. & Lwin, T. (2001). *J. Appl. Cryst.* **34**, 409–426.
Nath, N. K. & Nangia, A. (2011). *CrystEngComm*, **13**, 47–51.
Ong, L. R., Widjaja, E., Stanforth, R. & Garland, M. (2003). *J. Raman Spectrosc.* **34**, 282–289.
Oszlányi, G. & Sütő, A. (2004). *Acta Cryst.* A**60**, 134–141.
Patterson, A. L. (1935). *Z. Kristallgr.* **90**, 517–542.
Pawley, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.
Petříček, V., Dušek, M. & Palatinus, L. (2006). *JANA2006*. Institute of Physics, Prague, Czech Republic.
Putz, H. & Brandenburg, K. (2012). *Endeavour*. Crystal Impact GbR, Bonn, Germany.
Rietveld, H. M. (1969). *J. Appl. Cryst.* **2**, 65–71.
Scarlett, N. V. Y. & Madsen, I. C. (2006). *Powder Diffr.* **21**, 278–284.
Schreyer, M., Guo, L., Tjahjono, M. & Garland, M. (2011). *J. Appl. Cryst.* **44**, 17–24.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sin, S. Y., Widjaja, E., Yu, L. E. & Garland, M. (2003). *J. Raman Spectrosc.* **34**, 795–805.
Smith, J. H., Dann, S. E., Elsegood, M. R. J., Dale, S. H. & Blatchford, C. G. (2005). *Acta Cryst.* E**61**, o2499–o2501.
Thirunahari, S., Aitipamula, S., Chow, P. S. & Tan, R. B. (2010). *J. Pharm. Sci.* **99**, 2975–2990.
Visser, J. W. (1969). *J. Appl. Cryst.* **2**, 89–95.
Watanabe, S. (1981). *Pattern Recognit.* **13**, 381–387.
Werner, P.-E., Eriksson, L. & Westdahl, M. (1985). *J. Appl. Cryst.* **18**, 367–370.
Widjaja, E. & Garland, M. (2008). *Anal. Chem.* **80**, 729–733.
Widjaja, E., Li, C. & Garland, M. (2002). *Organometallics*, **21**, 1991–1997.
Widjaja, E. & Tan, W. J. (2008). *Appl. Spectrosc.* **62**, 889–894.