Structure determination of Fe(II) spin-crossover complexes from powder diffraction data with direct-space methods

Dova, E.

Citation for published version (APA):
Chapter 2

Techniques for structure determination from X-ray powder-diffraction data

2.1. Introduction

X-ray diffraction is one of the most important and powerful techniques for characterization and structure determination of crystalline solids. Single-crystal X-ray diffraction in particular is used routinely for crystal structure determination. However, not all crystalline solids can be prepared as single crystals of sufficient size and quality for single-crystal X-ray diffraction studies. Moreover, in some cases it is more convenient or desirable to study a material in a polycrystalline form, for example pharmaceuticals (optimal solubility properties), catalysts (increased surface area) and pigments (increased dispersibility). Therefore, it is essential to develop methods that enable the structural characterization of materials in this form and it is not surprising that the field of structure determination from powder diffraction data has become an active area of research in which new and/or improved methodologies and techniques are developed.

Although both single-crystal and powder diffraction data are based on the same structural information, in the latter case information is lost since the three-dimensional single-crystalline diffraction data is confined into a single dimension (intensity vs 2θ). The main consequence of this is the overlap of diffraction maxima (reflections) in the powder diffraction pattern, either accidental, especially in low-symmetry cases at higher angles, or systematic due to symmetry, especially in high-symmetry cases. Another consequence is that, in general, the final structural parameters obtained by powder diffraction data are not as accurate or precise as those determined by single-crystal diffraction data and it can not always be concluded reliably whether small deviations from standard bond lengths and angles are significant or not.

Crystal structure determination from diffraction data can be divided in general into the following stages: (1) unit-cell and space-group determination (indexing), (2) full pattern
decomposition/intensity extraction, (3) structure solution and (4) structure refinement. This chapter will discuss the first three subjects, and especially the several approaches that exist nowadays in structure solution from powder diffraction data, while the structure refinement procedure is discussed in the next chapter.

2.2. Unit Cell Determination – Indexing

The first step in a structure elucidation procedure is the determination of the lattice parameters (unit cell) and space group. In almost all structure determination techniques it is possible to enter the structure solution stage only if the correct unit cell has been determined. The information necessary to determine the unit cell concerns the accurate positions of the diffraction maxima and in most cases at least 20 peak positions are required. A significant zero-point error, ambiguous peak positions due to peak overlap, poor sample crystallinity or low instrumental resolution and the presence of an unknown impurity phase may impede or prohibit the determination of correct unit cell parameters. The most widely used programs for indexing powder diffraction data are ITO (Visser, 1969), TREOR (Werner et al., 1985) and DICVOL (Boultif & Louër, 1991). Also other less widely used programs exist, most of them being included together with the above-mentioned ones in the program package CRYSFIRE (Shirley, 2001).

An alternative manner for unit cell determination has been proposed by Kariuki et al. (1999) that uses a genetic algorithm in combination with whole-profile fitting. In this method, trial unit cells are generated using a genetic algorithm strategy and their quality is assessed with a whole-profile fitting procedure.

After the unit cell has been assigned correctly, the space group has to be determined taking into account the systematic absences. However, mainly because of peak overlap, the intensities and thus also the systematic absences may not always be inferred correctly so it is quite possible that the space group cannot be assigned uniquely. In such cases considerations of the density may be helpful but in the end all possible space groups should be considered.

2.3. Full Pattern Decomposition – Intensities Extraction

In structure solution, the aim is to obtain an (approximate) model of the crystal structure while initially no knowledge on the atomic arrangement in the unit cell is available. If this model is good enough, a refinement of the structure can be carried out successfully.

As it is known from single-crystal structure determination techniques, knowledge of the intensities of the reflections is essential for the structure solution since it contains information about the atomic arrangement in the unit cell. If these intensities would be known, methods developed for single-crystalline data could be applied to solve structures from powder diffraction data. However, peak overlap is often the main reason that the extracted individual reflection intensities may not be completely reliable. As a result, the application of methods that rely upon the correctness of the extracted intensities in the structure solution from powder data (direct
methods, Patterson methods, entropy maximization and likelihood ranking method) is often restricted to structures of a limited complexity.

The most common techniques to extract intensities are based on profile-fitting procedures and include a least-squares approach to fit the calculated profile to the experimental pattern by refining the parameters that characterize a powder diffraction pattern, like zero-point error, background parameters, unit cell and peak-shape parameters. Such methods work without reference to any structural model but, inevitably, the intensities determined from overlapping peaks by pattern decomposition will contain intrinsic uncertainties, especially when the difference in 2θ-positions of the peaks is smaller than half the full width at half maximum (FWHM) of the peaks.

The most widely used methods of fitting a powder diffraction pattern are the Pawley method (Pawley, 1981) and the Le Bail method (Le Bail et al., 1988). The first one treats intensities as individual variables in the least-squares fitting so that some of them may become negative during refinement, although this has no physical meaning. The second procedure is an iterative method that redistributes intensities starting from initial structure factors arbitrarily set to have the same value. A lot of work has been done to improve the extraction of intensities (Toraya, 1986; Jansen et al., 1992ab; Rodriguez-Carvajal, 1990; Altomare et al., 1995).

As will be discussed in 2.5, the direct-space techniques, an alternative approach for structure solution from powder diffraction data, do not rely on the availability of all individual reflection intensities, thus avoiding the peak overlap complication to a large extent, but still an analysis of the diffraction pattern is required in order to establish as reliable as possible the description of the peaks (accurate values of unit cell, zero-point offset, peak width, peak shape, background) to be used later in the calculation of the $R_{wp}$ for the trial structures.

2.4. Structure-solution Techniques

One of the earliest methods to solve crystal structures is the ‘trial and error’ method. This process involves the postulation of an atomic arrangement that conforms to the space group and the calculation of the structure factor amplitudes that this arrangement would give for certain reflections. If these amplitudes agree roughly with those observed then the atomic arrangement is probably correct, and one can continue the calculations for other reflections. However, if the calculations disagree with the observations, the atomic arrangement is probably wrong and discarded. In order to increase the probability to find the correct structure and because of the lack of certainty in this method, any hints possibly related to the atomic positions should be used. Thus, one should take into consideration spatial characteristics (e.g. atomic radii, atomic coordination, symmetry elements) and/or physical properties (e.g. piezoelectricity). The method has been applied for several inorganic and organic compounds by, for example, Taylor and Jackson (1928), Beevers and Schwartz (1935) and Robertson (1933). In organic structures, several relative positions of the atoms in the molecules or the complete stereochemistry is often known and it might be possible to construct complete structural models or fragments of the structure, which can be used in the same way as described above. Naturally, the main problem in the application of such methods is their time consumption, however, the development of the computer power in the recent years allowed for the reconsideration of their possibilities. Indeed,
the ‘trial and error’ method is the basic idea of the so-called direct-space methods developed in the recent years especially for structure solution from powder diffraction data. Such methods have been used in this thesis and they will be discussed separately in the next section.

As pointed out by Lipson and Cochran (1953), whatever the methods used for the structure determination, the proof of its correctness lies on the agreement between the calculated and the observed intensities. This leads to the idea of producing a structure from the observed intensities themselves, and several techniques have been developed, in which the atomic coordinates are being estimated from manipulation of the observed intensities as obtained from single-crystal diffraction data or as extracted from a powder diffraction pattern. The major examples of this class of techniques will be here discussed shortly. For more details and references on these techniques, the reader is referred to the books of Giacovazzo et al., (2002), Clegg et al., (2001), Schenk, (1991) and, Lipson and Cochran, (1953).

**Patterson methods:** The Patterson function $P(u) = \frac{1}{V} \sum |F_H|^2 \exp[-2\pi i \mathbf{H} \cdot \mathbf{u}]$ uses the observed structure factor amplitudes $|F_H|$ obtained from the reflection intensities $I_H$ that have been extracted in the pattern decomposition (Patterson, 1934). Each peak of the Patterson map $P(u)$ corresponds to a superposition of parallel interatomic vectors $(\mathbf{r}_i - \mathbf{r}_j)$ in the unit cell while its height is proportional to the sum of the products of the scattering powers of the atoms $i$ and $j$ in each of the parallel vectors. In principle, the atomic positions can be calculated from the set of interatomic vectors of a Patterson map but in practice only if the structure contains a small number of strong scatterers: the interatomic vectors involving these atoms dominate the Patterson map, allowing the positions of the strong scatterers to be readily deduced. However, if all non-hydrogen atoms have more or less the same scattering power, the Patterson map cannot be interpreted unambiguously. In case a structure (mainly) consists of a rigid body, a well-defined set of interatomic vectors appears in the Patterson map that can be used in positioning this rigid body. This procedure involves a rotational and a translational search in order to achieve the best fit of the calculated and observed Patterson map. A program for structure determination from powder diffraction data incorporating the above procedure is ROTSEARCH (Rius and Miravitlles, 1987).

**Direct methods:** Direct methods are based on the fact that certain items of direct-space information, e.g. that the electron density is always positive or that it is a superposition of spherically symmetric atoms of approximately the same shape, impose limitations on the reciprocal-space distribution, i.e. the magnitudes and the phases of the structure factors. By using relationships among the phases $\phi_H$, it is attempted to evaluate them directly from the measured diffraction intensities $I_H$. In the first steps of the structure determination with direct methods, as many phase relationships (triplets, quartets) as possible are collected and the origin is fixed by specifying the phases of a few suitable reflections numerically. In general, it is not possible to assign numerical phase values to all reflections just from the origin defining phases. However, it is possible to construct starting sets consisting of more reflections, which are capable to access the other phases through the use of phase relationships by means of a process called phase propagation or phase extension, through the multi-solution or the symbolic addition method. When a phase has been assigned to most of the strong reflections, the numerical values of the
unknowns are evaluated and using a Fourier summation an image of the structure is produced via the electron density equation (1.4).

Initially, direct-method programs developed for use in single-crystal determination procedure were used but later direct method procedures optimized for powder data have been developed, e.g. SIMPEL, (Jansen et al. 1992ab, 1993) and SIRPOW (Cascarano et al., 1992; Altomare et al., 1994).

**Maximum Entropy Method:** The Maximum Entropy Method (MEM) is a method to reconstruct the most probable and least biased probability distribution given a set of experimental data and it is applied in a wide range of scientific fields (Gull & Daniel, 1978; Buck & Macauley, 1992; Gilmore, 1996). In crystallography the maximum entropy criterion has been used to examine the crystallographic inversion problem (Wilkins et al., 1983), to generate high quality electron density maps (Livesey & Skilling, 1985; Sakata et al., 1990), as a means of partitioning the intensities of completely overlapping reflections in powder diffraction patterns and as an approach for direct phase determination (Narayan & Nityananda, 1982; Britten & Collins, 1982; Piro 1983; Bricogne, 1984). The method adopts a similar approach to structure determination as conventional direct methods. Both methods consider a crystal structure consisting of atoms with unknown position. These positions are considered as random variables having a uniform distribution in the asymmetric unit. By gradually removing this randomness the final crystal structure occurs. The method is able to handle groups of overlapping peaks enabling intensity information of these peaks to be used in the structure solution process. The program MICE (Gilmore and Nicholson, 1994) applies the MEM for structure solution from powder diffraction data.

### 2.5. Direct-Space Methods

In direct-space methods, as in the ‘trial and error’ methods mentioned in the previous section, trial crystal structures are generated in direct space independent of the experimental powder diffraction data. In this model-based approach the unit cell and space group should be known as well as the contents of the unit cell to a good approximation. The position and structural arrangement of the constituents are the subject of the investigation. The structure is defined by a fragment $X(j)$ comprising an appropriate collection of atoms within the asymmetric unit and it is defined by a set of variables $x(j)$, also called its degrees-of-freedom (DOF), that determines its location in the unit cell. Thus, a fragment can be defined by the fractional coordinates $(x, y, z)$ of the center-of-mass or of a predefined atom, its orientation by rotation angles $(\varphi, \psi, \kappa)$ with respect to a set of orthogonal axes and its intramolecular geometry by a set of $n$ variable torsion angles $(\tau_1, \tau_2, \ldots, \tau_n)$, see e.g. Andreev & Bruce (1998). It is more convenient to let these variables define the position and orientation of the structural fragment as a whole, rather than to consider the coordinates of every atom individually. It is possible to extend these concepts in case the structure consists of more than one structural fragment.

A structural fragment to be used in a direct-space method can be retrieved from a database or it can be constructed taken into account standard interatomic bond distances and angles for the type of molecules under study. Additionally, an energy minimization procedure may have to be carried out to find a suitable intramolecular configuration. In general, the complexity of direct-
space structure-solution calculations increases as the number of variables defining the structural fragment increases.

The correctness of each trial structure is assessed with a cost function, comparing the experimental powder diffraction data with the calculated pattern, either directly via the diffraction pattern or using the extracted intensities and/or clusters of intensities. A cost function or criterion-of-fit that is often being used is the weighted profile $R_{wp}$, as known from the Rietveld refinement. As the comparison is using the digitized pattern, the peak overlap is implicitly taken into account. Other definitions of cost functions such as $\chi^2$ (Andreev et al., 1997) or $R(X)$, based on extracted peak intensities (Chernyshev & Schenk, 1998; David et al., 1998), have also been used. The aim of a direct-space investigation is to find the trial structure with the lowest possible cost function.

Direct-space techniques have given a new prospective to determine the crystal structure of molecular compounds for which the success of traditional methods is relatively limited (Harris et al., 2001). The algorithms used in direct-space techniques range from exhaustive ones (Grid Search), in which a systematic search of the whole parameter space specified by the user (defined as grid) is carried out in predefined steps, to more sophisticated ones (Monte Carlo, Simulated Annealing, Genetic Algorithms etc), in which a selective search of the parameter space is performed on the basis of several criteria. Successful results have obtained using the above methods, to be discussed below in more detail.

2.5.1 Monte Carlo/Simulated Annealing Techniques

Monte Carlo and Simulated Annealing techniques (Metropolis et al., 1953; Allen & Tildesley, 1987; van Laarhoven & Aarts, 1987) are closely related and both involve the generation of a sequence of structures for consideration as plausible structure solutions. Monte Carlo techniques have been applied in structure determination from powder diffraction data since 1989 (Deem & Newsam, 1989; Newsam et al., 1992; Harris et al., 1994).

The first generated structure $X(j)$ is chosen by positioning the structural fragment randomly in the unit cell. Then, starting from structure $X(j)$, a series of trial structures $X(j+n; n=1,...)$ is generated with each structure, $X(j+n)$, being derived from the previous one, $X(j+n-1)$, by small random displacements in the set of variables $x(j)$ ($x, y, z, \phi, \kappa, \tau_1, \tau_2, ..., \tau_n$). After each change of variable the powder diffraction pattern of the trial structure is calculated and its $R_{wp}$, or any other cost function, is compared with that of the starting structure. If $Z = R_{wp}[X(j+n)] - R_{wp}[X(j+n-1)] \leq 0$ the trial structure is automatically accepted. If $Z > 0$, the trial structure is accepted with a probability of $\exp(-Z/S)$ and rejected with a probability of $[1 - \exp(-Z/S)]$ (using the Metropolis importance sampling technique), where $S$ is a scaling factor. So the new “starting” structure is then the $X(j+n+1)$ (if accepted) or $X(j+n)$ (if $X(j+n+1)$ was rejected). The scale factor $S$ is a parameter that expresses the strictness of acceptance of a trial structure; the higher $S$, the higher the probability that a trial structure with $Z > 0$ will be accepted. After an extensive range of structural space has been explored, the structure with the lowest $R_{wp}$ is selected to be used in the refinement.

In the Monte Carlo method, the parameter $S$ is fixed so a certain percentage of the trial structures will be accepted. This value might be altered during the Monte Carlo calculations in
order to allow exploration of a wider range of parameter space or a selected region in more
detail.

Starting from a single reference structure, Simulated Annealing techniques also use the
Monte Carlo algorithm to generate a series of structures, which will be accepted or rejected using
the Metropolis importance sampling, but the fundamental difference is the use of the parameter
S. In simulated annealing S is systematically decreased, like a decrease in temperature, according
to a certain annealing schedule. The annealing procedure starts from a ‘temperature’ S high
enough to allow the system to explore the entire conformational space. The temperature is
gradually lowered so that, in the later stages of the procedure, the local regions around a deep
minimum can be investigated.

A method proposed as an improvement of Simulated Annealing is Parallel Tempering
(Falcioni & Deem, 1999). Instead of a single reference structure in Simulated Annealing, the
idea of Parallel Tempering is to consider several reference structures in parallel, each evolving at
a constant simulation temperature. The temperatures are chosen such that the structure with the
highest temperature is able to search the complete conformational space, whereas the one with
the lowest temperature performs a local search for the best solution within the neighborhood of a
deep minimum. When a structure is selected for the next move, either a configurational move or
a swap move can be carried out. In the first case, a new trial structure is generated and assessed
in the same way as in the Simulated Annealing techniques. In the case of a swap move the
exchange of two structures at different temperatures is proposed. The proposed move is accepted
also in accordance with the Metropolis importance sampling. In this way the system may escape
from local minima.

A Monte Carlo technique combined with a potential energy function has been recently
proposed to work with low-resolution powder diffraction data as well as with compounds of
which the correct molecular structure or bonding pattern are not completely known (Brodski et
al, 2003). When a Monte Carlo move is realized, for this configuration the locally minimized $R_{wp}$
and a potential energy function $E$ (currently simple van der Waals potentials) are calculated
independently. The new configuration is accepted if either $E_{j+n} < E_{j+n}$ or $R_{wp(j+n+1)} < R_{wp(j+n)}$,
otherwise the old configuration is retained. The calculation of the potential energy function may
help to avoid chemically and physically unrealistic structures, which could have an acceptable
low $R_{wp}$ value. Overall in this Monte Carlo algorithm both $R_{wp}$ and the potential energy function
are expected to decrease, but not necessary simultaneously because a decrease of one cost
function can be accompanied by an increase of the other. By using two functions the system can
pass barriers in multidimensional space and escape from local minima on the $R_{wp}$/potential
energy hypersurface.

2.5.2. Genetic Algorithm

Genetic Algorithms (GA’s) (Holland, 1975; Goldberg, 1989; Cartwright; 1993) are
optimization techniques based on the principles of natural evolution and involve the evolutionary
operations of crossover, mutation and natural selection.

The Genetic-Algorithm approach is based on the evolution of a set of trial crystal
structures (population). The initial population consists of a number of randomly generated
structures (parents) that may evolve to subsequent generations by evolutionary operations. Each member \(X(j)\) of the population is defined by a unique set of variables \(x(j)\). The initial structures are ranked according to a cost function (\(R_{wp}\) or \(R(X)\) or others, depending on the method). The top ranking members of the initial population may propagate into the next generation via a natural selection procedure.

A certain percentage of the trial structures is allowed to participate in a crossover procedure in which some of the parameters defining the structures may be swapped (the precise procedure depends on the implementation). The probability that a structure propagates and its participation in crossover operations depend on its ranking position. The higher the position of a trial structure (parent) in the ranking, the higher the possibility to be selected for crossover. Every two parents produce two new structures (the offspring).

Offspring and parents form an intermediate population that goes through a mutation procedure in which every parameter has a certain (predefined) possibility to mutate. The mutation rate can be adapted during the calculations in order to get a population with a higher or lower diversity. Another way of enriching diversity is the inclusion of a certain number of randomly generated structures. At the end, after ranking all the structures (offspring, mutants, new members, old members selected by natural selection), a new population is selected with a size that is equal to the initial population. The evolutionary cycle is repeated for a specified number of generations or until convergence is reached.

The application of the GA in structure determination from powder diffraction data was realized independently by two research groups: the group of Harris and co-workers (Kariuki et al., 1997; Harris et al., 1998ab; Lanning et al., 2000; Kariuki et al., 1999; Tedesco et al., 2000; Turner et al., 2000) and the group of David and co-workers (Shankland et al., 1997 & 1998; Csoka et al., 1998). Another implementation of a GA has been realized recently in the Laboratory of Crystallography of the University of Amsterdam (see 2.7.2). All three approaches differ in the definition and handling of the cost function and in some aspects of the way the genetic algorithm is implemented.

For example, in an implementation of Turner et al (2000), each structure generated during the calculation was subjected to local minimization of \(R_{wp}\). These minimized structures were used for further evolutionary calculations. This approach was shown to improve the efficiency of the calculations both in terms of the necessary number of generations to achieve convergence and in reproducibility of finding the correct solution.

In another implementation (Lanning et al, 2000) a combined normalized energy/\(R_{wp}\) cost function was used. At high energy, the energy guided the calculation towards energetically plausible structures while at lower energies an increasing importance was given to the \(R_{wp}\).

2.5.3. Grid Search

Grid Search (GS), the most straightforward direct-space method, is an exhaustive search procedure in which the structural fragment is moved systematically throughout the unit cell. All possible packing arrangements of the fragment are assessed systematically by changing the values of the variables of the set \(x(j)\) using predefined grid increments. The more degrees of freedom the structural fragment has, the more time consuming the search is.
The implementation of this technique as incorporated in the program MRIA (Zlokazov & Chernyshev, 1992) will be discussed in 2.7.1. Another program performing an exhaustive grid search is P-RISCON (Masciocchi, 1994), that also uses intensities of single and/or sets of overlapping reflections but does not take texture into account.

2.6. Some remarks on the structure solution methods

The reliability of the techniques involving extraction of intensities, is increased by the use of high-resolution data (e.g. synchrotron) as the peak overlap is in general lower than in conventional X-ray data. For direct-space techniques, which use the $R_{wp}$ as the cost function, synchrotron data are not necessarily advantageous over laboratory data (Harris et al., 2001).

The complexity of the structure-determination problem in the methods described in 2.4 depends on the number of atoms to be located in the asymmetric unit whereas in direct-space methods the complexity is defined by the set of parameters $x(j)$ (DOF) of the structure. Prior structural knowledge, e.g. structural units of well-defined geometry or extended framework structures consisting of well-defined building blocks, is essential in case of direct-space methods because it reduces the DOF and it is being used in the calculations. Thus, large rigid molecules can be solved in general easier with direct space techniques than smaller but flexible ones.

Recently, hybrid approaches have been developed that incorporate features of the global optimization techniques into direct methods (Altomare et al., 1998, Burla et al., 2002).

In general, direct-space algorithms require more computational time than the other structure-solution methods described in 2.4 and it is recommended to run the algorithm several times for different starting structures: obtaining a specific solution several times is a strong indication of its correctness.

2.7. Structure solution methods used in the current work

2.7.1. Grid Search

The method as described in 2.5.3 has been implemented in a local version (Driessen, unpublished) of the program suite MRIA (Zlokazov and Chernyshev, 1992). A full pattern decomposition (FPD) procedure is carried out in the beginning with which integrated intensities and/or clusters of intensities with value $X_{obs} = \sum m_i |F_{o,i}|^2 = s \cdot \sum m_i y_i |F_{c,j}|^2 = s \cdot X_{calc}$ can be extracted. The summation is over the overlapping reflections, $m_i$ is the multiplicity of the $i^{th}$ reflection, $y_i$ is a texture correction multiplier and $s$ an overall scale factor. The cost function $R(X)$ (Chernyshev and Schenk, 1998) used to assess the correctness of the calculated structures is defined as $R(X) = \sum_j |X_{obs}(j) - s \cdot X_{calc}(j)| / \sum_j X_{obs}(j)$. An advantage of the $R(X)$ is that the overlapping reflections can be included in a natural way in the calculation of the cost function. Texture can be taken into account during calculation if the direction of preferred orientation is known or by applying the spherical harmonics method in which the predefinition of the orientation of the preferred orientation is not necessary.
Quite often in case of flexible molecules, the computational time required makes the method less efficient. However, in some GS problems it is possible to split the total set of variables in smaller sets, each next set having a smaller (expected) influence on the cost function. By carrying out sequential sub-searches for each set it is not necessary to consider all variables simultaneously and unnecessary calculations can be avoided (Dova et al., 2001).

2.7.2. Genetic Algorithm

As a part of the program suite MRIA, a version of the GA method has been developed at the Laboratory of Crystallography (University of Amsterdam, The Netherlands) (Driessen, unpublished).

The program is able to handle four independent structural fragments. The parameter space includes the translational (in fractional coordinates), rotational as well as torsion angle parameters. Optionally, the parameter space can be limited automatically to a smaller one (by 0.37) around a current best solution. The cost function $R(X)$ used is the same as for Grid Search in MRIA.

The mutation rate (the probability that a structure will mutate) is adapted automatically during the calculations in the range $[0.0005, 0.25]$ in order to get a population with a higher or lower diversity and its default starting value is 0.005. The crossover rate (default 0.95) is the probability that two structures will crossover. A number of the best solution is kept (default 10 structures) in every new generation. Several GA tests showed that multiple cycles with a small population (150-250) and few generations (200 –300) is to be preferred over a single cycle with a larger population and many generations (with respect to both calculation time and effectiveness).

Parallel Genetic Algorithm (pGA) is a variation of the method described above, in which several GA cycles are processed in parallel, examining a smaller parameter space (box). A box is created around the solution of a cycle (typical sizes are 1.5 Å, 10° and 30° for translational, rotational and torsion parameters respectively) and its size is dynamically varied in order to search a smaller or larger area. The pGA allows the search within five boxes in parallel, in addition to the search in the user defined parameter space.

2.7.3. Parallel Tempering

Parallel tempering as incorporated in Powder Solve (Engel et al., 1999) of the Materials Studio package (Accelrys, 2001) has been used for the structure solution of several structures in this work. In general, the default values of the parameters guiding the algorithm have been used and an overview is given below.

The highest temperature is calculated automatically. The program calculates the standard variation of the $R_{wp}$ factor using a random sequence of trial structures. The highest temperature is set to 1.5 the value of the standard variation. Typically, the highest temperatures are in the order of 0.02, corresponding to acceptance ratios larger than 0.8. The lowest temperature is calculated automatically and is set to 0.2 times the highest temperature.
The step width is the average variation of the value of a degree of freedom during a single Monte Carlo step (not to be confused with the fixed step width in the exhaustive Grid Search procedure). Step widths are random and have values in the range 0 to 1 and for translational degrees of freedom they are measured in fractions of a lattice vector, for rotational and torsional degrees of freedom they are fractions of 360 degrees. By default, the program adapts the step width automatically, such that the acceptance ratio is approximately 0.5.

During a Powder Solve run, suitable trial structures are selected and subjected to a local Rietveld minimization, which ensures that the local minimum is found accurately.

The program also allows searching for preferred orientation of crystallites in a powder sample using simulated annealing. The direction of the preferred orientation vector and the values of the preferred orientation parameters (the preferred orientation degrees of freedom) can be varied in addition to the structural degrees of freedom.

References

Accelrys (2001). Materials Studio. Accelrys Inc., 6985 Scranton Road, San Diego, CA 92121-3752, USA


Beevers, C. A. and Schwartz, C. M. (1935). Z. Kristallogr. 91, 157-


