Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 1 May 2002
Accepted 9 July 2002

Determination and refinement of disordered crystal structures using evolutionary algorithms in combination with Monte Carlo methods

Thomas Weber[a]* and Hans-Beat Bürgi[b]

[a]Laboratory of Crystallography, ETH Zentrum, CH-8092 Zurich, Switzerland, and [b]University of Berne, Laboratory of Chemical and Mineralogical Crystallography, Freiestrasse 3, CH-3012 Berne, Switzerland. Correspondence e-mail: weber@kristall.erdw.ethz.ch

An evolutionary algorithm called 'differential evolution' is combined with Monte Carlo simulation to determine and optimize models of disordered crystal structures. Requirements for successfully finding the parameters describing disorder from diffuse scattering data are discussed and the algorithm is applied to resolving the racemic and associated displacive disorder of the host substructure in a perhydrotriphenylene inclusion compound. Refinement resulted in a very good visual agreement between observed and calculated intensities and in a relatively low value of $R_{\rm diffuse} = 0.148$ (3). The computations for determining and refining the structure took 29 d with five to ten workstations running in parallel. Analysis of the progress of the structure determination shows that the essential information can be obtained within a few hours. Limits of the technique and strategies to optimize the procedure are discussed.

1. Introduction

Disorder in crystals implies departure from perfect translational order and is associated with diffuse scattering. A quantitative characterization of such disorder from diffuse scattering proceeds in three steps: (i) deriving as much information as possible from the average structure about the types of disorder (occupational, positional, displacive, orientational, conformational *etc.*); (ii) qualitative interpretation of the diffuse scattering and quantitative modelling of faults including their spatial distribution; (iii) refinement of the parameters defining the model. Before fast computers and Fourier transform algorithms were available, analysis of diffuse scattering was restricted to relatively simple frequently one-dimensional problems, for which the observed intensities could be described in terms of one or a few analytical equations with a small number of parameters (see *e.g.* Guinier, 1963; Rosshirt *et al.*, 1985; Weber *et al.*, 2000). With the advent of more efficient hard- and software, more complicated disordered structures have been tackled, relatively large model crystals simulated and their Fourier transforms compared with experiment (Proffen & Welberry, 1997; Welberry *et al.*, 1998, 2001).

Current methods of analysing disordered structures combine Monte Carlo (MC) simulation for model building with different ways to refine the model parameters. The most developed technique is direct MC simulation (Proffen & Welberry, 1998). The initial model crystal is set up using disorder information from the average structure and neglecting all correlation of disorder between neighbouring unit cells. The lattice energy of the crystal is calculated with interatomic or intermolecular potentials and minimized by randomly replacing or displacing atoms or molecules. A new conformation is accepted if $\Delta E = E_{\rm new} - E_{\rm old} < 0$. Otherwise the old conformation is restored with a certain probability. The Fourier transform of the equilibrated model crystal is calculated and compared with the experimental data. If the fit is unsatisfactory, the parameters are adjusted manually and the MC simulation is repeated. The performance of the refinement is usually low, but the technique has the advantage that the disordered structure may be described by relatively few parameters, which may have direct chemical significance.

The method was improved by introducing automatic refinement of the model parameters based on a least-squares algorithm (Welberry *et al.*, 1998). Although applied successfully to a number of disordered systems (Mayo *et al.*, 1999; Welberry, 2000; Welberry *et al.*, 2001), the method does not overcome the problem intrinsic to all non-linear least-squares refinements: if the initial parameter values are not sufficiently close to the global minimum, optimization may be trapped in a local minimum, *i.e.* refinement may converge while still maintaining noticeable differences between observed and calculated quantities. In order to ascertain a high probability of reaching the global minimum, several sets of parameters need to be tested and the computations may become very time consuming.

While the search for the best parameters is done manually with the direct MC technique, the automatic MC technique is a manual search with an automated local optimization. In this paper, we describe a new method for analysing disordered crystal structures, which provides a more general search of optimal MC parameters using an evolutionary algorithm called 'differential evolution'.

2. Evolutionary algorithms and differential evolution

Evolutionary algorithms are optimization techniques that mimic the Darwinian principle of natural evolution in the computer. They are especially powerful for solving problems that are too complex to be described analytically (see *e.g.* Michalewicz, 1996). Evolutionary techniques have been applied to several crystallographic problems. Shankland *et al.* (1997), Kariuki *et al.* (1997) and Harris *et al.* (1998) used them to solve crystal structures directly from powder data. Landree *et al.* (1997) combined evolutionary strategies with direct methods for the determination of surface structures. Wormington *et al.* (1999) and Ulyanenkov *et al.* (2000) have refined thin-film structures. In macromolecular crystallography, Chang & Lewis (1997) proposed a molecular replacement strategy based on an evolutionary algorithm. Knorr & Mädler (1999) were the first to apply evolutionary strategies to the study of disorder in crystals using disorder information exclusively from the average structure.

Like most other optimization techniques, evolutionary algorithms require a parametric model of the problem. Using the language of evolution, a parameter of the model is called a *gene*, a vector of genes is called a *chromosome*. The parameter values specify the *genotype* of a virtual *individual*. Its appearance to its environment is its *phenotype*. At the beginning of an evolutionary process, a starting set of individuals (*population*) with different genotypes is created. New individuals (*children*) are created by recombination (*crossover*) and/or mutation (*variation*) of genes within the initial generation (*parents*). The *fitness* of parents and children is determined by comparing the corresponding phenotypes with predefined requirements. Only the fittest individuals from the combined population of parents and children survive and form the next generation of parents. This procedure is repeated for a specified number of generations or until the fitness of the population shows no further improvement.

There are many mechanisms for recombination and mutation and many criteria for selecting the fittest individuals. For an overview of algorithms, the reader is referred to the literature (*e.g.* Goldberg, 1989; Michalewicz, 1996). Here we restrict ourselves to the variant called differential evolution (DE) as developed by Price & Storn (1997). It is said to be easy to use and very efficient in finding the global minimum for problems that may be parameterized in terms of real numbers. DE is special in the way it handles recombination and mutation (Fig. 1). Each individual of a parent generation serves as one parent for an offspring. This parent is called the target individual. The second parent is not taken directly from the parent generation but is computed from three randomly chosen individuals a, b, c in the parent generation. Their chromosomes $\mathbf{p}_a$, $\mathbf{p}_b$ and $\mathbf{p}_c$ are combined according to

$$\mathbf{p}'_c = \mathbf{p}_c + f_m(\mathbf{p}_a - \mathbf{p}_b). \quad (1)$$

The mutation constant f_m is a parameter of the algorithm. The chromosome of a child is obtained by modifying the genes of the target individual: one randomly selected gene is replaced by the corresponding gene from $\mathbf{p}'_c$ to ensure at least one difference between the child and the target individual. The child inherits the remaining genes from $\mathbf{p}'_c$ with a probability given by the crossover constant f_r, another parameter of the algorithm. In the new generation of parents, the child replaces the target individual if its fitness is higher. Otherwise the target individual survives. This procedure is repeated until a stop criterion is reached.

The degree of mutation in DE scales with the diversity of the parent population. To ensure efficiency of DE, the initial values of the genes should be taken at random from a range of values that covers all reasonable models. As long as the diversity is large, the difference vector $\mathbf{p}_a - \mathbf{p}_b$ in (1) has a high probability of containing large components and the range of parameter search is large as well. Note that this algorithm

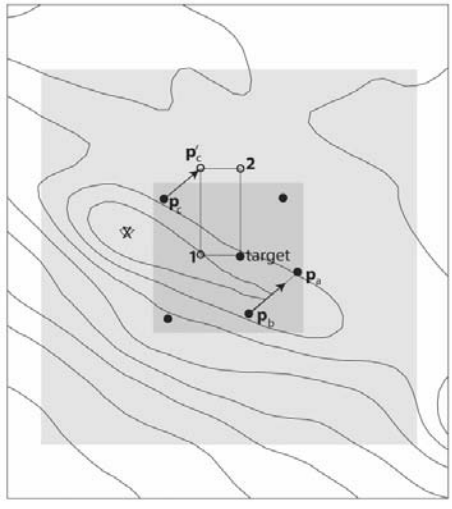

Figure 1
Illustration of the DE algorithm in two-dimensional parameter space. The global minimum of the example is indicated by X. The filled circles show six individuals of a parent population. The intermediate child $\mathbf{p}'_c$ results from its parent $\mathbf{p}_c$ using (1), whereby the arrows represent the term $f_m(\mathbf{p}_a - \mathbf{p}_b)$ from (1) with $f_m = 0.75$. Possible children to be compared with the target individual are indicated by open circles. The probability of selecting child **1**, **2** or $\mathbf{p}'_c$ as a new trial is governed by the parameter f_r. Child **1** (**2**) is obtained if the gene represented by the horizontal (vertical) axis is taken from $\mathbf{p}'_c$ and the other is taken from the target, while $\mathbf{p}'_c$ is selected if none of the target genes are inherited. Note that $\mathbf{p}'_c$ is selected as a trial individual if $f_r = 1$, while **1** or **2** are taken if $f_r = 0$, because child and target chromosomes must differ in at least one gene. The dark grey square indicates the region of individuals of the parent generation. The fact that child **2** and $\mathbf{p}'_c$ are outside the region covered by the parent generation illustrates the ability of the differential evolution algorithm to extend the search beyond the parameter range originally covered. The light grey plus dark grey squares give the volume accessible to the children.

allows genes of a child to lie outside the initial region of parameter values.

If several local minima populated with several individuals are found, $\mathbf{p}_a - \mathbf{p}_b$ becomes small if $\mathbf{p}_a$ and $\mathbf{p}_b$ happen to be from the same minimum; otherwise the difference remains large. In such an intermediate stage of the optimization, small difference vectors imply a local search in the neighbourhood of a minimum, while large difference vectors imply continuation of the global search. Finally, when all individuals are concentrated in only one minimum, hopefully the global one, all difference vectors are relatively small and the optimization has again local character.

3. The use of differential evolution in combination with Monte Carlo algorithms for the analysis of diffuse scattering

According to Michalewicz, 'a genetic algorithm (as any evolution program) for a particular problem must have the following five components:

- a genetic representation for potential solutions to the problem,
- a way to create an initial population of potential solutions,
- an evolution function that plays the role of the environment, rating solutions in terms of their "fitness",
- genetic operators that alter the composition of children,
- values for various parameters that the genetic algorithm uses (population size, probabilities of applying genetic operators *etc.*).' (Michalewicz, 1996, pp. 17*ff.*)

In the following, these requirements are adapted to the investigation of diffuse intensities using the MC technique. The use of constraints and restraints, the handling of experimental data and the parallelization of DE are also considered.

For the sake of simplicity, the MC algorithm is described for a crystal consisting of rigid orientationally ordered molecules. Treatment of ionic and metallic compounds as well as of solids consisting of flexible and orientationally disordered molecules requires only slight modifications to the procedure described below. Without loss of generality, the present discussion is further restricted to harmonic pair potentials and disregards correlations between model parameters. Special attention is paid to the economic use of computing time because refinement of disordered structures using MC techniques is computationally extremely expensive. The cost is even higher for evolutionary strategies because the search for the best solution is not directed by a least-squares algorithm but requires a smart trial-and-error technique.

Genetic representation: The genetic representation of a problem implies an encoding of its physical content into a model (genotype) and a decoding of the model into experimentally observable quantities (phenotype). Each model parameter (gene) has two attributes: firstly, its *meaning* in the context of the model and, secondly, its numerical *value*. The crucial step is the definition of meaning, *i.e.* of the relevant interactions between molecules in the crystal. A description of disordered crystals with atom–atom potential as used for the prediction of average structures would be far too time consuming for the simulation of disordered crystals. Instead, the intermolecular interactions are described by a simplified set of interaction parameters, which is both adequate and of a manageable size. At present, the problem of choosing adequate parameters is best solved with the help of experience, chemical intuition and trial-and-error methods. For the present purpose, a MC model is defined in terms of pair interaction energies $E_{ij}(x) = J_{ij} + k_{ij}(d_{0,ij} - x)^2$ between neighbouring molecules i and j. J_{ij} is an interaction constant independent of the exact positions of molecules i and j, k_{ij} is a Hook's law force constant, $d_{0,ij}$ is the equilibrium distance between molecules i and j and $(d_{0,ij} - x)$ is the deviation from the equilibrium distance. The genotype is decoded into the phenotype by executing the MC algorithm. This leads to the model crystal and the corresponding calculated diffraction intensities. Because of the simplified representation of intermolecular interactions, the genes have no direct physical meaning, only the phenotype has, since it is related to experimental observations.

In principle, the question of meaning could also be treated with evolutionary algorithms. Several techniques are based on populations for which genes in the same position of a chromosome have different meaning depending on the individual. Even the number of genes need not be the same for all individuals (Michalewicz, 1996). However, a discussion of such algorithms is beyond the scope of this paper. In DE, which is designed to optimize real numbered values of genes, the meaning and the number of genes are the same for all individuals.

Initial population: At the beginning of an evolutionary procedure, a population of individuals with different and random genotypes is defined by selecting the values of the model parameters from reasonable ranges of starting values. In as much as the physical meaning of the genes is limited, such ranges cannot be defined easily from physical arguments alone, but may have to be determined by trial and error or by qualitative and semi-quantitative analysis of the diffuse intensities. An example is described by Weber *et al.* (2002).

The fitness function: An X-ray diffraction experiment measures reciprocal-space properties of a phenotype. Therefore, the fitness of the genotype is defined as the agreement between observed and calculated diffuse intensities, usually expressed as $\chi^2 = \sum w(\Delta I)^2$ or as

$$R = \left(\sum w(\Delta I)^2 \Big/ \sum w I_{\rm obs}^2 \right)^{1/2}.$$

The summations include all reciprocal-space points used in the analysis and w is a weighting factor for an individual data point. The difference ΔI is calculated as $I_{\rm obs} - (b + sI_{\rm calc})$, where b denotes the background and s is a scale factor. In order to keep the number of parameters small and to speed up their refinement, it is recommended that the scale factor and background correction be obtained as described by Proffen & Welberry (1997).

For automatic MC refinement, weights are usually chosen as $w = 1/I_{\rm obs} \sim 1/\sigma^2(I_{\rm obs})$ (Welberry *et al.*, 1998; Mayo *et al.*, 1999; Welberry, 2000; Welberry *et al.*, 2001). This choice takes into

account that high intensities are more affected by statistical noise in the experimental data as well as in the calculated intensities (see below). A disadvantage of this choice is that contributions to the R value from narrow diffraction features, which usually occupy a small volume of reciprocal space and are more intense than broad ones, are downweighted whereas the weaker broad features, *i.e.* disorder phenomena of short correlation length, are overemphasized. Further, with $w = 1/I_{obs}$, systematic errors in the determination of the scattering background strongly affect the goodness of fit. For these reasons, unit weights were used in this study.[1]

The χ^2 and R values reflect not only the quality of the experimental data and the weights but also the quality of the model and the respective calculated intensities. Both of these depend not only on the meaning of the parameters but also on the statistical properties of the model. Simulations always result in model crystals, which are small compared with the real crystal. Their characteristic probabilities fluctuate about a mean value and are propagated to their Fourier transforms and calculated diffraction patterns, which may be much more noisy than the experimental one.

Welberry & Proffen (1998) have described a technique to reduce the influence of statistical fluctuations in simulated structures on diffraction patterns. A number of small lots of equal size are selected from a relatively large model crystal. Their size has to be larger than the maximum correlation lengths represented by the diffuse data chosen for interpretation. The lots are distributed randomly and may overlap. The Fourier transform is calculated for each single lot and the intensities from all lots are summed up incoherently. The variance of the calculated diffuse data decreases as the number of lots increases.

The presence of noise in I_{calc} implies that the χ^2 hypersurface is not sharp as it is in the refinement of an average structure, where it has a unique value for a given set of parameters. Statistical fluctuations in the structures of the lots produce statistical fluctuations in the χ^2 values even for identical parameter values (apparent fitness). As a consequence, an individual of high apparent but low intrinsic fitness may outdo an individual of low apparent but high intrinsic fitness, especially when the simulated crystal and the number of lots are small, *e.g.* at the beginning of a refinement (see below). With an increasing number of lots, the variance of the calculated fitness of a given individual decreases and consequently the probability that an individual of low intrinsic fitness outdoes an individual of higher intrinsic fitness decreases as well. Thus, the fuzziness of the χ^2 surface is not only a problem, it also provides a welcome mechanism for escaping from local minima and sampling extended areas of parameter space in the early stages of refinement. The number of lots for calculating intensities may be seen as an analogue to the model temperature in the well known technique of simulated annealing (Kirkpatrick *et al.*, 1983). A small number of lots corresponds to a high model temperature.

In the early stages of a refinement when the model is still poor, the influence on χ^2 of statistical noise tends to be low compared with the influence of inappropriate values of the genes. One may therefore start with a low number of lots. If statistical noise in the calculated intensities begins to interfere with the refinement, the lot number must be increased. If this does not decrease χ^2 significantly, the evolution has reached the global or a local minimum. The structural model should then be checked for consistency with known structural data. In case of inconsistencies, a new DE refinement with a different starting population should be performed or a different genetic representation of the model should be attempted. Note that a population, or a major part of it, may remain in a local minimum for a long time. Thus, evolution should be stopped only if results are satisfying or if further improvements exceed the computational resources.

Artefacts in the calculated Fourier transform due to finite lot size (series-termination effects) can be minimized by taking into account the number of unit cells n_x, n_y and n_z along the corresponding directions. The Fourier transform of a lot is the sum of a periodic Bragg-like scattering based on the corresponding lattice function

$$L(hkl) = \frac{\sin(\pi h n_x)}{\sin(\pi n h)} \frac{\sin(\pi k n_y)}{\sin(\pi k)} \frac{\sin(\pi l n_z)}{\sin(\pi l)}$$

and of a continuous diffuse part. The lattice function has nodes at non-integral reciprocal-lattice points $h = m_h/n_x$, $k = m_k/n_y$ or $l = m_l/n_z$ with integer m_h, m_k and m_l; it shows crystal truncation ripples in between. Contributions to the diffuse signal from such ripples may be avoided by subtracting incoherently from the total diffraction pattern the intensities corresponding to the *average* structure of the lots. Because a single and an average lot have the same dimension, their truncation ripples have the same nodal properties and truncation effects will disappear upon subtraction. In this study, diffuse intensities are calculated only at the positions where the lattice function and thus the contributions from the crystal truncation ripples are zero. With this method, the average intensity need not be calculated (Neder & Proffen, 1999).

Data binning and data selection: In general, the computational resources are insufficient to refine a complete diffuse data set. The first problem is discretization of the continuous diffuse intensities by binning the data to a grid. The grid size should be such that a minimum of data points gives a sufficient representation of the diffuse diffraction profiles of interest. Even after an optimized binning, the number of data points is usually still far too large. It is further reduced by selecting those areas of reciprocal space containing most information on the disorder. Such reduction must be done very carefully to avoid bias in the results.

Genetic operators and parameters of DE: Mutation and recombination operations in DE have been described in §2. The parameters required for DE are the mutation constant f_m, the crossover constant f_r and the population size. Their best values depend on the problem at hand and need to be found

[1] In principle, a weight matrix with non-zero off-diagonal elements based on the variance–covariance matrix of the experimental observations would be more appropriate because diffuse diffraction data are continuous.

empirically (Price & Storn, 1997). Because DE combined with MC techniques is very time consuming, it is impractical to test different values of f_m and f_r and some guidelines for their choice may be useful. The crossover constant f_r determines the probability that a gene of the child's chromosome comes from the parent vector $\mathbf{p}'_c$ [equation (1)]. With $f_r = 1$, the child's chromosome is identical with $\mathbf{p}'_c$ and, in general, completely different from all chromosomes in the parent generation. With $f_r = 0$, all but one, randomly selected, genes of a child come from the target individual. The genetic material present in a population will change only slowly even if children replace their parents frequently. Satisfying results have been obtained for crossover constants f_r in the whole range between 0 and 1 (Price & Storn, 1997).

An interesting special case illustrates these considerations: populations with all parameters linearly correlated, *i.e.* distributed along a one-dimensional valley of the n-dimensional χ^2 hypersurface. In this case, the difference vectors $\mathbf{p}_a - \mathbf{p}_b$ tend to point along the valley and thus the endpoint of the intermediate vector $\mathbf{p}'_c$ in n-dimensional space is situated in the valley as is the endpoint of vector $\mathbf{p}_c$. For $f_r = 1$, the child is $\mathbf{p}'_c$ and thus the search is restricted to the bottom of the valley. For $f_r = 0$, the search is along only one of n coordinates which in general does not coincide with the direction of the valley. The first scenario may drastically speed up the refinement if the bottom of the valley leads towards the global minimum, while the second case may help to escape from the valley by 'tunneling' through the surrounding χ^2 walls.

Wormington *et al.* (1999) used a slightly modified version of DE. Instead of the randomly selected chromosome $\mathbf{p}_c$ of (1), they used the chromosome of the best individual. This restricts the search space to the environment of the best individual rather than to the environment of all individuals. While this may lead to faster convergence, it also increases the risk of getting trapped in a local minimum.

The mutation constant f_m controls the volume covered by the search. A large value of f_m may be useful to extend the volume of search space into the neighbourhood of the initial volume, while a small value of f_m restricts the search and may be useful for final optimization in the region of the global minimum. Reasonable values of the mutation constant f_m have been reported to be in the range $0.4 \leq f_m \leq 1$.

The population size has to ensure sufficient sampling of parameter space while minimizing the substantial computational effort involved in the calculation of fitness. Price & Storn (1997) reported good results with a population size that exceeds the number of refined parameters by a factor of 5 to 10.

Restraints and constraints representing prior knowledge about a structure may be included in DE calculation to speed it up and improve the structure model. Given the limited physical meaning of most MC parameters, restraints and constraints on the limiting values of genes should be applied with great care unless they are known from previous experience. A less error-prone strategy is to apply restraints and constraints to the real-space phenotypes, *i.e.* to the model crystals from the MC simulation. Such models must conform to the *average* structure obtained from Bragg refinements. They should not contradict general chemical information as well as qualitative and semi-quantitative conclusions from preliminary qualitative investigations of the diffuse diffraction pattern. If they do, their fitness must be reduced. A 'death penalty' may be imposed on strongly violating phenotypes independently of their fitness value. The latter avoids the calculation of the expensive Fourier transformation and may therefore speed up the refinement, especially in the early stages of evolution when completely unreasonable individuals occur frequently.

Parallelization: For every genotype of a generation, a structure is simulated and subdivided into single lots; these are Fourier transformed and the fitness of the individual is calculated. The computations of fitness are easily parallelized by distributing them to different, independent, computers connected *via* a network or to different processors on a multiprocessor computer. In addition, the Fourier transformation of individual lots from the same model crystal may be similarly optimized. The best strategy depends on available computer resources and should minimize the involvement of the network and the need for synchronization of the processes. In this work, a scheme of distributed computing was used.

4. Application to a real world problem

In this section, DE is applied to resolving the host disorder in a perhydrotriphenylene (PHTP) inclusion compound with 1-(4-nitrophenyl)piperazine (NPP) as a guest molecule. A detailed discussion of the MC part of the model and the resulting disordered structure will be given elsewhere (Weber *et al.*, 2002).

4.1. Nature of disorder to be analysed

The host lattice of the $PHTP_5 \cdot NPP$ crystal consists of stacks of triangular PHTP molecules which form a honeycomb-like tunnel architecture parallel to the orthorhombic c axis (König *et al.*, 1997) (Fig. 2). The space group is $Cmc2_1$ with $a = 15.023$, $b = 23.198$, $c = 4.73$ Å at 100 K. The guest molecules are enclosed in the tunnels with their long axis parallel to the tunnels. Each stack of PHTP molecules is surrounded by three other PHTP stacks, one along the b axis and two along the positive and negative a directions. Neighbouring PHTP stacks differ in their z position by $c/2$, *i.e.* by half a stacking distance. The average structure of the host shows occupational disorder with half an *R*-PHTP molecule superimposed on half an *S*-PHTP molecule. The aim of our DE calculations was to find the distribution of *R*- and *S*-PHTP molecules and to analyse the diffraction pattern for displacive disorder associated with occupational *R*/*S* disorder.

4.2. Experimental observations and data selection

Diffraction data were recorded at $T = 120$ K with synchrotron radiation at the Swiss Norwegian Beamline at ESRF in Grenoble, France, using an image-plate detector. They were transformed to reciprocal-space coordinates using

Table 1
Sections from experimental data used in the refinement and *R* values corresponding to the best individual obtained by DE.

Layers *a*3–*f*3 cover a somewhat shorter range along **c*** than *a*2–*f*2 because of strong diffuse guest scattering in the layer *hk*3.2; layers *A*2, *B*2, *A*3 and *B*3 are shown in Fig. 3(*c*).

Layer	*h*	*k*	*l*	*nh*	*nk*	*nl*	Individual *R* value
*A*2	5–10	0–3	2	51	31	1	0.16
*B*2	2–5	7–11	2	31	41	1	0.14
*A*3	5–10	0–3	3	51	31	1	0.14
*B*3	2–5	7–11	3	31	41	1	0.14
*a*2	6.5	0–2	1.8–2.2	1	21	21	0.12
*b*2	6.8	0–2	1.8–2.2	1	21	21	0.09
*c*2	7.2	0–2	1.8–2.2	1	21	21	0.10
*d*2	7.5	0–2	1.8–2.2	1	21	21	0.09
*e*2	2.5	7–11	1.8–2.2	1	41	21	0.14
*f*2	3.5	7–11	1.8–2.2	1	41	21	0.12
*a*3	6.5	0–2	2.8–3.16	1	21	19	0.12
*b*3	6.8	0–2	2.8–3.16	1	21	19	0.16
*c*3	7.2	0–2	2.8–3.16	1	21	19	0.21
*d*3	7.5	0–2	2.8–3.16	1	21	19	0.12
*e*3	2.5	7–11	2.8–3.16	1	41	19	0.20
*f*3	3.5	7–11	2.8–3.16	1	41	19	0.11

the program *XCAVATE* (Estermann & Steurer, 1998; Scheidegger *et al.*, 2000). Further experimental detail is given by Weber *et al.* (2001). The information on the distribution of *R*- and *S*-PHTP molecules and an associated displacive disorder is concentrated in a subset of pseudo-hexagonal diffuse columns parallel to **c*** and indexed by hexagonal indices (*HK*). Cross sections through the columns at integer *l* show relatively broad modulations of irregular shape (Fig. 3). Three general features of the modulations are worth noting: (i) the shape of the diffuse features are very similar in the *hk*1 and *hk*3 layers; (ii) high diffuse intensities in the *hk*1 and *hk*3 layers tend to be accompanied by low intensities at the same *hk* positions in the *hk*2 layer and *vice versa*; and (iii) diffuse intensities around the Bragg positions are highly asymmetric along **a***. The symmetry of the diffuse scattering from the columns appears to be orthorhombic. The modulations described for *hk*1 – *hk*3 are hardly observable in the *hk*0 layer (Weber *et al.*, 2001), indicating that this layer exhibits practically no information on *R*/*S* disorder. This lack of information is due to the fact that *R*- and *S*-PHTP molecules cannot be distinguished in projection down the *c* axis (Fig. 2).

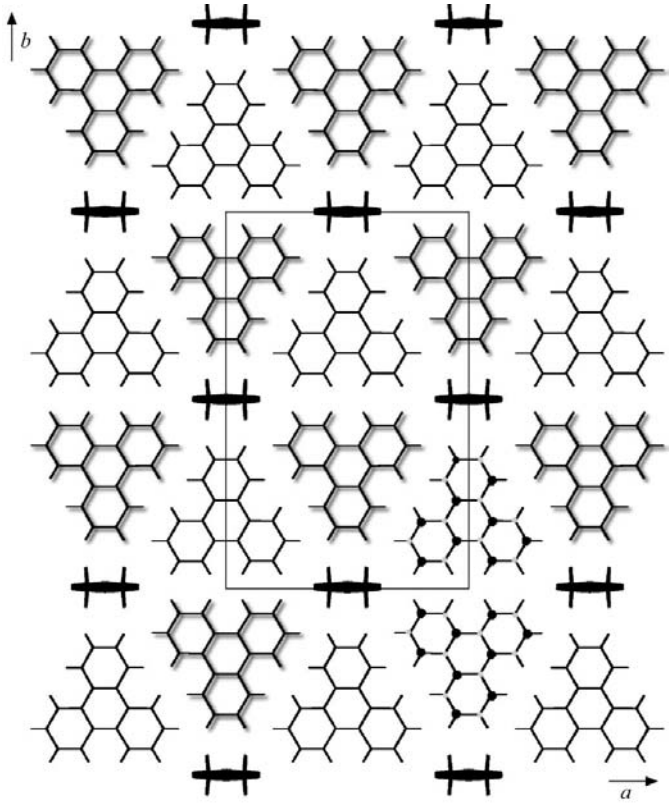

Figure 2
The structure of the $PHTP_5$·NPP inclusion compound viewed along the *c* axis. The wavy shape of an *R*- and an *S*-PHTP molecule is shown in the bottom right by black and grey carbon atoms above and below the molecular plane, respectively. Different *z* positions of the molecules as a whole are indicated by the presence ($z = 0.75$) or absence ($z = 0.25$) of a shadow.

Diffuse intensities are high at positions close to integer *l* and weak in between (see Fig. 4 in Weber *et al.*, 2001). The profiles near integer *l* are slightly asymmetric along **c***, broader at $|l| + |\Delta l|$ than at $|l| - |\Delta l|$ (Fig. 4*a*, this work). Their width varies as a function of *h* and *k*. The profiles of the diffuse data included in the refinement are broader by at least a factor of 10 than the profiles of those Bragg reflections that are weak enough not to become saturated during the 60 s exposure. Since profiles of Bragg reflections represent an upper limit of the resolution function, the observed diffuse profiles can be calculated as a convolution of experimentally observed Bragg profiles with calculated diffuse profiles. Assuming that the shapes of Bragg reflections and of diffuse features are approximately Gaussian, the convolution produces another Gaussian with variance $\sigma^2_{\rm exp} = \sigma^2_{\rm diffuse} + \sigma^2_{\rm Bragg}$. Because $\sigma^2_{\rm diffuse} \approx 100\sigma^2_{\rm Bragg}$, the effect of instrumental resolution on the widths of experimental profiles is less than 1% and has been neglected.

Data selection: A minimum of data has been selected that represents as much information as possible on the *R*/*S* host disorder and as little as possible about any other kind of

disorder. Fortunately, the required information is concentrated in the narrow (*HK*) columns. The (10) and (01) columns were selected because they are stronger than the (20) columns and, unlike the (20) columns, have been recorded fully (Fig. 3). This choice includes a system of streak-like scattering unrelated to racemic disorder of the host (Weber *et al.*, 2001), but

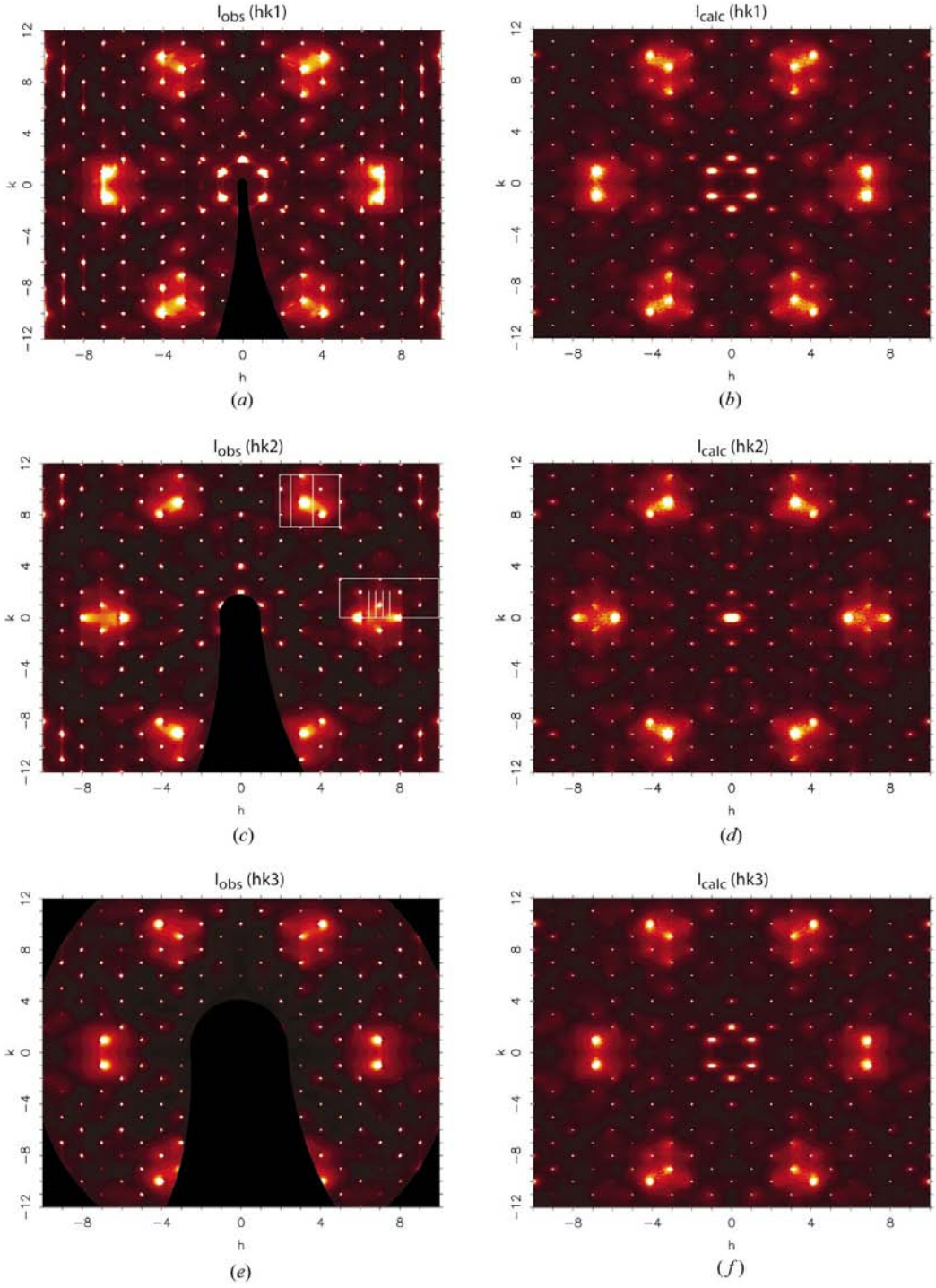

Figure 3
Observed and calculated intensities in the *hk*1, *hk*2 and *hk*3 layers. The regions of the *hk*2 layer, which form the experimental basis of the structure analysis are marked in (*c*). The rectangles show the selected areas within the $\mathbf{a}^*\mathbf{b}^*$ plane (*A*2 middle, *B*2 top), while the lines are traces from sections parallel to the $\mathbf{b}^*\mathbf{c}^*$ plane. Data from the *hk*3 layer were taken in corresponding positions (Table 1). Calculated intensities represent the fittest individual in the 220th generation of the reference calculation and are based on 150 lots from a model crystal consisting of 90 × 90 × 100 cells. Although no experimental data from the *hk*1 layer were used in the refinement, the agreement between observed and calculated intensities is good. Streak-like scattering parallel to $\mathbf{b}^*$, which can be observed at 701 and close to 98*x* (*x* = 1, 2, 3) positions, is not related to *R*/*S* disorder and cannot be reproduced by the current model. Smeared-out and additional reflections in the *hk*1 layer are due to twinning and to an additional phase of hexagonal symmetry (Weber *et al.*, 2001).

superimposed on all columns and potentially affecting the analysis of *R*/*S* disorder. To minimize such interference, the *hk*0 layer has been excluded from analysis because it shows strong streak intensities but contains practically no information on *R*/*S* disorder. The *hk*1 layer has been excluded for two reasons: first, streaking is still noticeable and, second, the information on *R*/*S* disorder is hardly different from that in the *hk*3 layer, which is not affected by streaking. This leaves diffuse intensities within and close to the *hk*2 and *hk*3 layers for interpretation.

Because of the finite and variable width of the profiles parallel to $\mathbf{c}^*$ (see Fig. 4 in Weber *et al.*, 2001), a refinement of sections taken exclusively from *hkx* layers at integer values of *x* would disregard information on correlations along the stacking direction. Therefore, several *xkl* sections with non-integer *x* were also considered (Table 1). Both the *hkx* and *xkl* sections were re-binned into a grid with step widths of $\Delta h = \Delta k = 0.1$ and $\Delta l = 0.02$. The grid size was chosen such that (almost) every one-dimensional scan along $\mathbf{c}^*$ includes at least four data points with intensities higher than the half-maximum of the feature. It guarantees a maximum of information about the shape of the features parallel to $\mathbf{c}^*$ and sufficient sampling of the background. In each section, positions of Bragg reflections as well as the eight bins surrounding the position of a Bragg reflection were marked as unobserved. This rather generous suppression of experimental data is necessary because all Bragg reflections in the regions of interest are heavily overexposed and their tails cover an area larger than the size of a single bin. The total number of observed data is 12344 in 16 sections.[2]

4.3. Definition of the differential evolution and Monte Carlo models

Model crystals and their genetic representation: Model crystals were constructed *via* MC growth simulations. PHTP molecules were placed in an **ab** seed layer as in the average structure but with random distribution of chirality. PHTP molecules were then added along **c**. Their position and chirality were determined with the assumption that an added PHTP molecule interacts with only four neighbouring PHTP molecules, *i.e.* with the top molecules in the growing stack and in the three neighbouring stacks. The probability for adding an *R* or an *S* molecule was assumed to follow a Boltzmann distribution with $p(S) = 1/\{1 + \exp[(E(S) - E(R))/kT]\}$, $p(R) = 1 - p(S)$ with *k* the Boltzmann constant and *T* the model temperature. The energy $E(S)$ corresponding to adding an *S* molecule to the stack is expressed as:

$$E(S) = \tfrac{1}{2}\sum_{i=1}^{4}(k_i\{[z(S) - z_i] - d_i\}^2 + J_i\sigma_S\sigma_i), \qquad (2)$$

z_i is the *z* coordinate of one of the four neighbouring molecules *i*, $z(S)$ is the equilibrium position of the molecule being

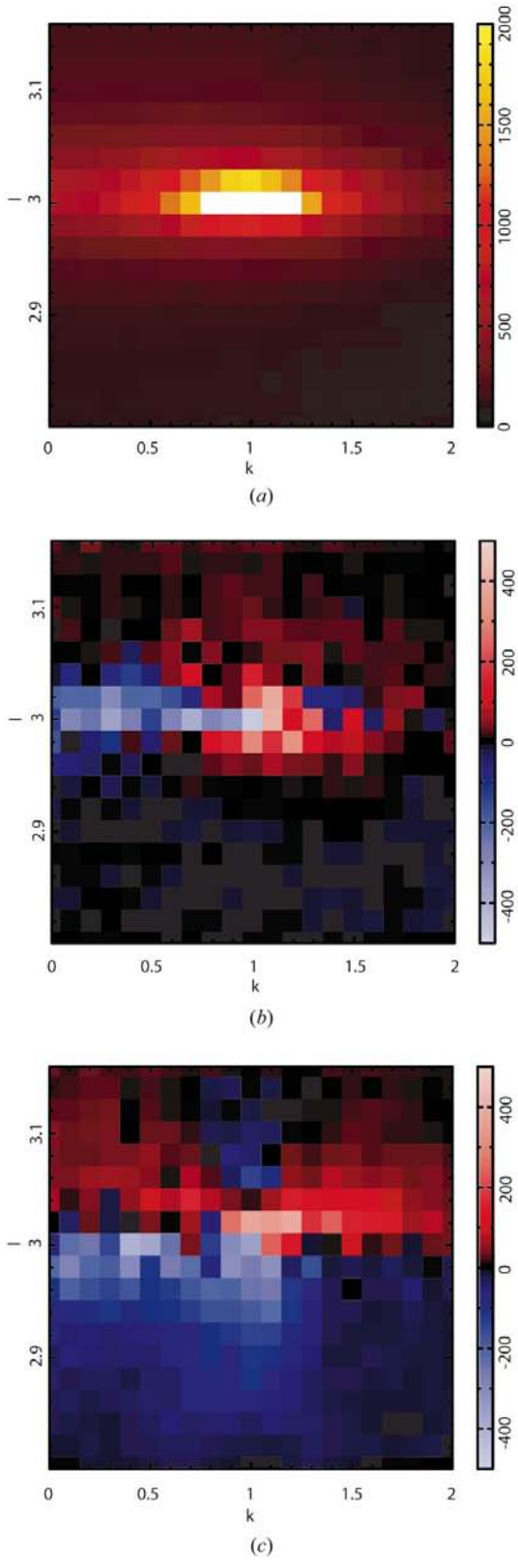

Figure 4
(*a*) Observed intensities in the 6.8*kl* layer (section *b*3, Table 1) and corresponding differences $I_{obs} - I_{calc}$ (*b*) from reference refinement and (*c*) from test refinement *F*. Calculated intensities used for maps (*b*) and (*c*) were obtained from the best individual in the pertinent final generation. The asymmetry of the profiles along $\mathbf{c}^*$ is better reproduced by the reference refinement than by test refinement *F*.

[2] Supplementary data for this paper, including observed intensities, are available from the IUCr electronic archives (Reference: SH0158). Services for accessing these data are given at the back of the journal.

added and is determined by four springs with spring constants k_i and equilibrium distances d_i. This position is

$$z(S) = \sum k_i(z_i + d_i) \Big/ \sum k_i.$$

The Ising-type parameters J_i represent pair interaction energies between neighbouring molecules and are independent of distance; σ is 1 for an R molecule and -1 for an S molecule. The calculations of $E(R)$ and $z(R)$ are analogous. A random number between 0 and 1 chosen from a uniform distribution is compared with $p(S)$. An S molecule is added at position $z(S)$ if the random number is smaller than $p(S)$. Otherwise, an R molecule is added at the position $z(R)$. The x and y coordinates of added molecules are assumed to be independent of R/S disorder and were taken from the average structure.

A semi-quantitative analysis of the diffuse intensities, symmetry considerations and chemical arguments suggested seven parameters, *i.e.* seven genes, for the description of the growth model: three pair interaction energies J_a, J_b and J_c and three force constants k_a, k_b and k_c along corresponding crystal directions and independent of the chirality of the interacting molecules; one parameter d_{RS} accounting for deviations from the average $\Delta z = 0.5$ if PHTP molecules in neighbouring stacks along the a axis are heterochiral. The equilibrium value of Δz for homochiral contacts along **a** and **b** was set to $\Delta z = 0.5$, that for contacts along **c** to $0.99|\mathbf{c}|$ for a homochiral and $1.09|\mathbf{c}|$ for a heterochiral contact. These values were derived from two considerations: first, homo- and heterochiral van der Waals contacts along **c** are expected to differ by about $0.1|\mathbf{c}|$ ($\sim$0.5 Å) owing to differences in non-bonded H$\cdots$H distances and, second, preliminary analysis has shown that the relative frequency of homochiral and heterochiral contacts is about 10:1; kT was kept constant at 1.

At the beginning of a DE calculation, the model crystals were grown over 600 layers along the c axis (= 300 unit cells) on a seed consisting of $50 \times 50 \times 1$ unit cells along **a**, **b** and **c**, respectively. Only the final 200 layers (100 unit cells) were used for the selection of lots and calculation of fitness to minimize artefacts propagating from the seed layer.

Initial population: The population size was 40 individuals. Using trial and error and a MC feedback mechanism as implemented in the program package *DISCUS* (Proffen & Neder, 1997), it was found that reasonable starting values of the spring force constants are in the range [0–100]. Starting values of the parameters J_a and J_b were taken from the interval $[0 \ldots 1]$, those of J_c from $[-1 \ldots 0]$. J_c must be negative to reproduce the positive correlations along **c**. Starting values of d_{RS} were taken from the range $[-0.05 \ldots -0.01]$ (in fractional z units, corresponding to $[-0.24 \ldots -0.047$ Å$]$). The selection of these parameter ranges is discussed in more detail elsewhere (Weber *et al.*, 2002).

Genetic operators and parameters of DE: The chromosome $\mathbf{p}_c$ in (1) was defined to be the best individual from the current population. The crossover constant f_r was set to 0.5 and the mutation constant f_m to 0.7.

Restraints: The following restraints were enforced: population of R- and S-PHTP molecules in the model crystal between 20 and 80% (expected value from the average structure is 50%); average translation from a new layer to the next-nearest layer below between 0.66 and 1.5 fractional units (expected value is 1); mean square deviation of the rigid molecules from their average position smaller than 0.9 Å^2 (expected value from atomic displacement parameters of the average structure is < 0.05 Å^2); positive correlations between homochiral molecules along **c** (diffuse intensities are highly concentrated at integer l). Failure to comply with any of these restrictions resulted in the death penalty. The restrictions are very generous but turned out to be very useful, particularly in the early stages of the refinement. Since the death penalty was applied before calculation of the Fourier transform, this procedure improved the efficiency of the refinement significantly.

Fitness function: The fitness of an individual is given by the χ^2 value based on intensities and unit weights. Initially, intensities were calculated from 20 lots. This number was later increased to 40, 80 and 150 (see below). The size of a single lot was $10 \times 10 \times 50$ unit cells along **a**, **b** and **c**, respectively. Note that the lot size matches the grid of the binned data. Fourier transformation takes into account four different kinds of PHTP molecules: R- and S-PHTP molecules, each in two different orientations related by a rotation of 180° about **c** (Fig. 2). A lot is separated into four sublattices, each containing only those points that are occupied by the respective kind of molecule. Using the convolution theorem, the Fourier transform of each sublattice is obtained by multiplying the Fourier transform of the sublattice calculated from the molecular coordinates obtained by MC growth with the corresponding molecular form factor calculated from the atomic coordinates reported by König *et al.* (1997). The results from the four sublattices are summed up coherently. The use of molecular form factors allows the inclusion of the hydrogen atoms of PHTP molecules without significant extra cost and accelerates the Fourier transformation enormously relative to a calculation based on atomic positions. The Fourier transforms of the sublattices and of the molecular form factors were calculated with the program package *DISCUS* (Proffen & Neder, 1997). The background and the scale factor were set to be the same for all reciprocal-space sections.

Computations: Generation of the model crystals and computation of their fitness function was distributed to as many as ten processors. A 'master computer', which is connected to 'slave computers' *via* a network, coordinates the computations and controls the progress of DE. At the beginning of the process, the master connects to a free slave *via* remote control (program *telnet*) and starts a program installed locally on the slave and supplies it with the chromosome of an individual. The slave then performs the crystal-growth simulation and the subsequent Fourier transformation. The master repeats this process until all available computers are busy. A slave that has finished its job connects to the master *via* file transfer protocol *ftp* and sends a file containing the resulting χ^2 value of the individual and, if requested, further information about the growth simulation and the corresponding Fourier transform. In addition, the slave asks

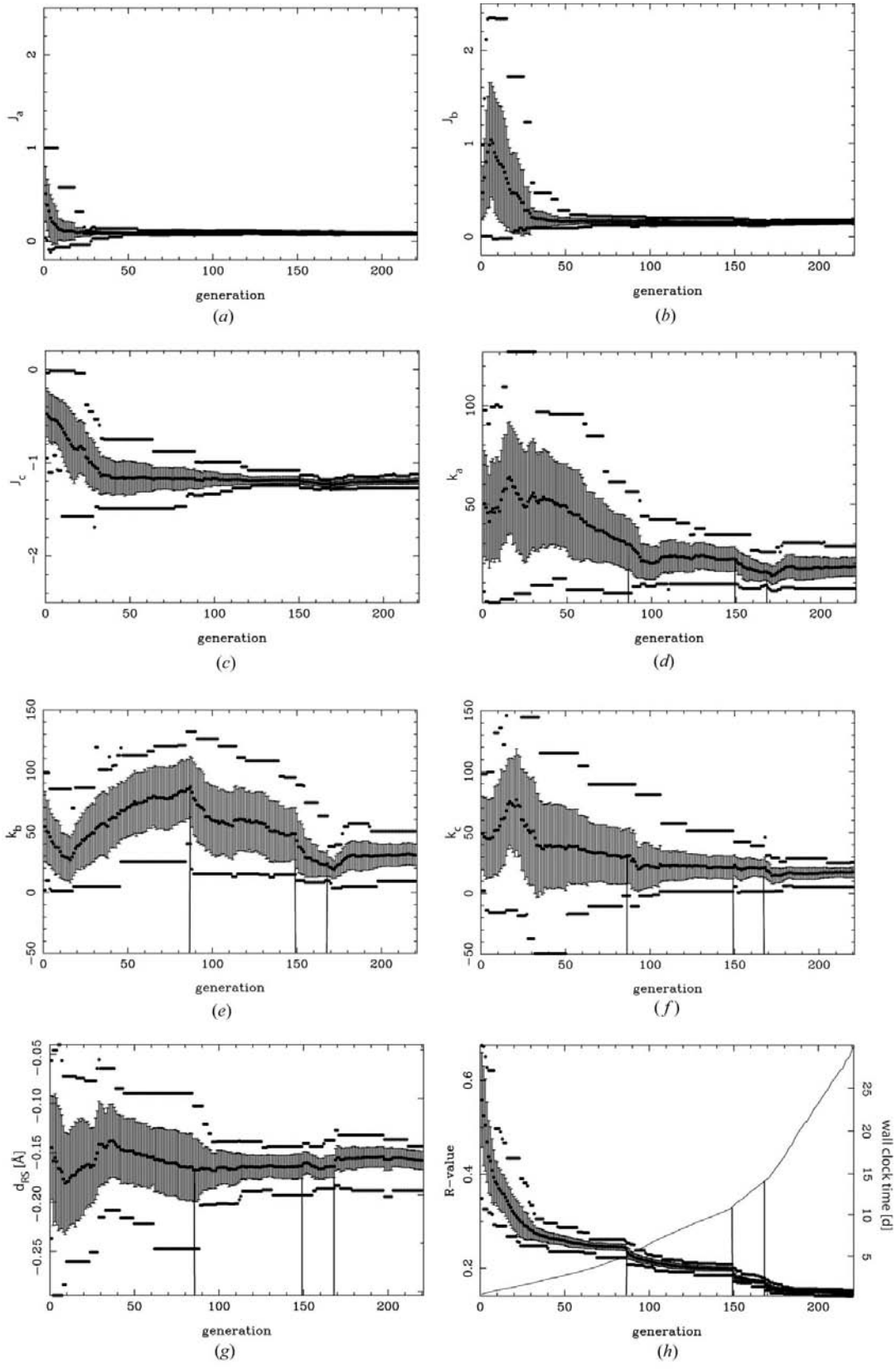

Figure 5
Evolution of the model parameters and of the unweighted *R* value. The bullets show minimum, average and maximum values and the error bars indicate r.m.s. deviations from the population average in each generation. The vertical bars indicate the generation at which the number of lots was doubled. In (*h*), the elapsed wall clock time is shown on the right.

the master *via telnet* to read the results and to supply the chromosome of a new individual. The master program keeps track of the jobs in progress, those that are finished and those that are in the waiting queue. When the number of submitted but incomplete jobs becomes smaller than the number of slaves, the master randomly selects a job from the 'in progress list' and resubmits it to a free slave. This redundancy avoids delays due to a slave that fails to report results for reasons that are beyond the control of the master, *e.g.* network errors. Although very simple, this redundancy procedure turned out to be very robust against communication errors. Finally, when the calculation of all χ^2 values of one generation is finished, the master stops the remaining but now redundant jobs, calculates the genes of the next generation and starts another round of calculating χ^2 values. The list of available slave computers is updated after every generation allowing computers to be added or removed dynamically depending on availability.

A Pentium MMX, 233 MHz computer running under Linux was used as the master computer. In addition, the following machines were available as slaves:

- six slots on a Sun SPARC workstation cluster running under the operating system Solaris 2.6;
- one Silicon Graphics Indigo 2 workstation running under IRIX;
- one Pentium III, 800 Mhz personal computer running under Linux.

Note that this rather heterogeneous collection of computers could be combined without any problems. Note also that the Pentium III computer was located at the University of Würzburg, Germany, *i.e.* several hundred kilometres from the master computer in Berne, Switzerland; it was included in the procedure without any problems. Depending on availability, it was possible to run five to ten jobs simultaneously. Given the heterogeneity of machines and the fluctuations in computing power, it is neither possible nor meaningful to report CPU time used. Instead, wall clock time is given. This provides a sufficient basis for crude comparisons of the time requirements for the different stages of the refinement.

4.4. Progress and results of a full calculation

The evolution of R values and parameters during the DE calculation is shown in Fig. 5. The R value decreased steeply in the first 25 generations. No significant improvement was found after about 80 generations. Visual inspection of the calculated diffraction patterns showed that most of the differences between observed and calculated intensities can be assigned to noise in $I_{\rm calc}$ because of the relatively small number of 20 lots (each with $10 \times 10 \times 50$ unit cells). After increasing the number of lots to 40 in generation 86, the R values improved significantly. Further improvement was obtained on increasing the number of lots to 80 and 150 after 149 and 168 generations, respectively. After 220 generations, the refinement was stopped because R values remained constant and a further increase of the number of lots was not practical with the available computer resources. The size of the model crystal from which the lots were taken has been increased in parallel with the number of lots such that its volume was not smaller than the sum of all lot volumes. The size was $50 \times 50 \times 100$ cells for 20 and 40 lots, $70 \times 70 \times 100$ cells for 80 lots and $90 \times 90 \times 100$ cells for 150 lots. The size of a single lot was always the same, $10 \times 10 \times 50$ cells.

The model parameters showed three kinds of evolutionary behaviour: J_a, J_b and J_c converged very quickly (Figs. 5*a*–*c*). After about 30 generations, they had almost reached their final values. The genes J_a and J_b showed a particularly steep decrease of genetic diversity. For parameters k_a, k_c and d_{RS}, diversity increased in the early stages of the evolution (Figs. 5*d*,*f*,*g*); after about 20 generations, the drift of the parameter average changed its sign towards smaller values but genetic diversity remained high. These parameters began to converge only after parameters J_a, J_b and J_c were close to their final values. The average of k_c was almost constant after generation 33; in subsequent generations, its diversity decreased slightly (Fig. 5*f*). Note that none of the parameters except k_b changed its average value significantly after generation 100, *i.e.* after six days. With further evolution, the diversity of the genes became smaller and the average values decreased slightly. A completely different behaviour was observed for the gene k_b (Fig. 5*e*). Its diversity and its average value started to converge only after increasing the number of lots to 80 in generation 149 and to 150 in generation 168. The evolution of all parameters illustrates an important feature of DE: the genetic diversity of a parameter tends to remain high or to increase as long as a parameter has not found its final value and as long as the noise in $I_{\rm calc}$ is high or both. This behaviour allows individuals to escape from local minima. The total time required for the DE calculations was substantial. With 20 lots, the calculation of the fitness for one generation took about one hour. Computation times doubled with each increase of the number of lots. In the last generations, progress was slowed down to no more than two or three generations per day. In total, the refinement took 29 d for 220 generations (Fig. 5*h*).

The R values of the 40 surviving individuals are between 0.141 and 0.153. The R values of single sections of data for the best individual are between 0.09 and 0.21 (Table 1). The agreement is illustrated in Fig. 3 for the layers *hk*2 and *hk*3 as well as for the layer *hk*1, which was not included in the refinement. Representative close ups of $I_{\rm obs}$, $I_{\rm calc}$ and $I_{\rm calc} - I_{\rm obs}$ from section *B*2 are displayed in Fig. 6. They show that even with 150 lots the noise in $I_{\rm calc}$ is much larger than in $I_{\rm obs}$. However, not all difference intensities are random. The negative differences shown in Fig. 6(*c*) encompass and connect the regions of the Bragg reflections 392, 3,11,2, 282 and 482. Fig. 4(*c*) shows that the asymmetry of the diffuse profiles along $\mathbf{c}^*$ is reproduced with high accuracy, although some systematic deviations seem to remain along $\mathbf{b}^*$. Comparison of the systematic deviations with the random noise in Fig. 6(*c*) indicates that the magnitude of the systematic error is about the same as that of random noise. This implies that systematic and statistical errors contribute to the R values about equally.

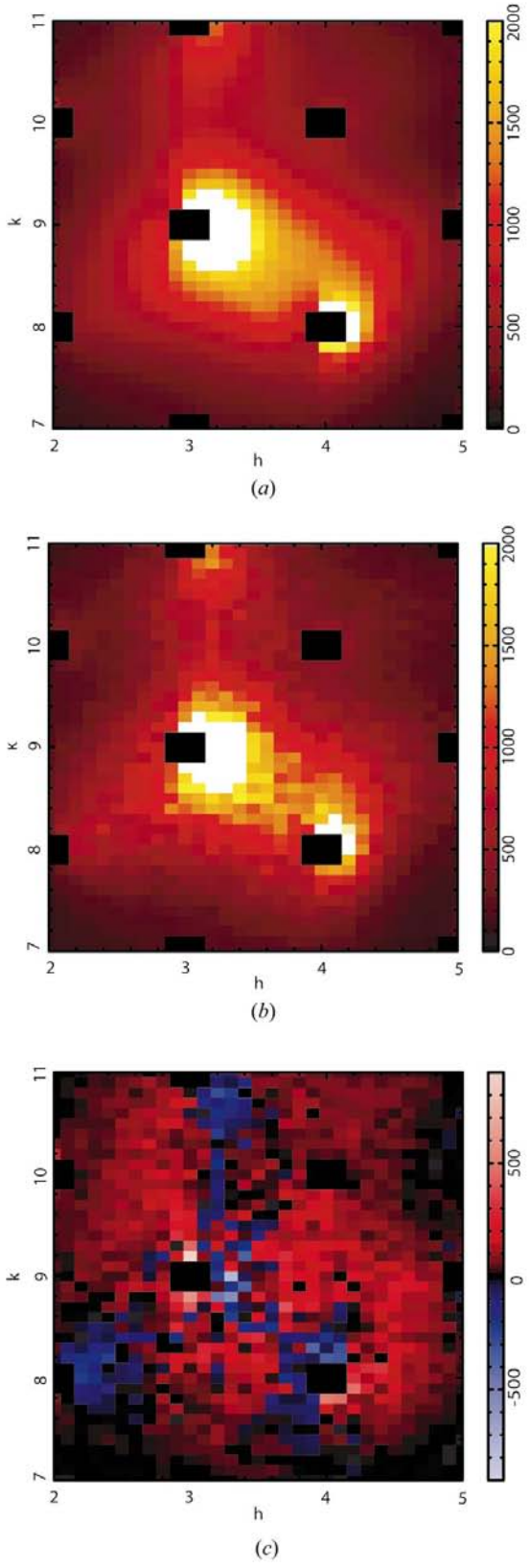

Figure 6
(*a*) A small part of the observed intensities, (*b*) calculated intensities and (*c*) corresponding differences $I_{obs} - I_{calc}$ in the *hk*2 layer. Calculated intensities used for maps (*b*) and (*c*) are obtained from the fittest individual in the 220th generation of the reference calculation using 150 lots from a model crystal with 90 × 90 × 100 cells.

The systematic differences between observed and calculated intensities may be due to: (i) insufficient flexibility of the disorder model; (ii) a DE refinement that got stuck in a local minimum; (iii) a DE refinement that has reached the area of the global minimum but was stopped before complete convergence; or (iv) errors arising from the processing of the experimental data. Computational limitations prevented testing these hypotheses. From the relatively low R value of $R_{diffuse} = 0.148$ (3), we conclude that refinement ended close to the global minimum and that the results are reliable.

4.5. Test calculations

Although the results described in the preceding section were satisfactory, the DE procedure was tested for its scope, *e.g.* its potential to find the global minimum if reasonable starting ranges for the parameter values are not known in advance and the influence of the number of lots and of the refinement parameters f_m and f_r on the performance of the algorithm. The available computational resources imposed the following limitations on these tests: refinements were not always run to the point where further qualitative or quantitative improvement could be excluded; they were stopped and considered successful when the values of the genes covered the initial range of the full calculation, they were stopped and considered not successful if the genes diverged away from the results found in the full calculation (in the latter case, one must check whether the results represent another, possibly better, solution to the problem). All test calculations were performed with the same population size of 40 individuals. It must be stressed that these criteria neither exclude nor guarantee that continuation of the refinement will lead to the results found in the full calculation. However, provided the genes reach the volume defining the initial values in the full calculation, there is a good chance that further refinement ends up in the same minimum as in the full calculation.

The following refinement strategies using an inflated start parameter volume were tried out:

A: the same as in the full calculation except for larger ranges of starting values. The range of initial values of k_a, k_b and k_c was increased by a factor of ten to [0 . . . 1000]. For J_a and J_b, the initial values were taken from the range [−10 . . . 10], for J_c from the range [−20 . . . 0], an increase by a factor of 20. The limits for the initial values of d_{RS} were set to −0.3 and 0.3. These changes correspond to an increase of the volume of the initial search space by a factor of about 10^8! As in the full calculation, the number of lots was doubled when χ^2 stopped decreasing.

B: like *A* with the mutation constant f_m set to 0.9.

C: like *A*, but the size of the model crystal reduced to 20 × 20 × 200 cells along **a**, **b** and **c**, respectively, and Fourier transform calculated from a single lot taken from the top 200 layers.

D: like *C* but with the chromosome $\mathbf{p}_c$ randomly selected rather than taken as the fittest of the current population.

E: like *C* but with weights set to $1/I_{obs}$.

Table 2
Comparison of the final results from the reference calculation with intermediate and final results from test calculations *F* and *G*.

For further details see text. The computing power available for the reference calculation was about twice that for calculations *F* and *G*.

	Reference calculation	Test calculation *F* after		Test calculation *G* after	
	Final result after 220 generations (29 d)	90 generations (30 h)	312 generations (45 d)	60 generations (14 h)	209 generations (70 h)
J_a	0.082 (4)	0.10 (2)	0.077 (4)	0.11 (3)	0.09 (2)
J_b	0.156 (6)	0.14 (2)	0.159 (6)	0.23 (7)	0.15 (2)
J_c	−1.19 (3)	−1.4 (2)	−1.28 (4)	−1.3 (2)	−1.3 (1)
k_a	18 (5)	3 (1)	3.1 (7)	57 (13)	34 (11)
k_b	30 (9)	1.7 (6)	6.0 (7)	103 (36)	75 (22)
k_c	17 (4)	−0.9 (5)	−3 (1)	32 (18)	21 (13)
d_{RS}	−0.032 (2)	−0.022 (5)	−0.035 (2)	−0.030 (6)	−0.038 (6)

The calculations were stopped after 1.5 to 4.5 d, corresponding to about 50 generations in *A* and *B* and 150 generations in *C*, *D* and *E*. Note that, because the number of lots was generally smaller in calculations *C*, *D* and *E*, a larger number of generations could be calculated.

None of the refinements converged to the results obtained in the full calculation (the reference solution). Inspection and comparison of calculated diffraction patterns clearly showed the results of the test calculations to be inferior to those of the full calculation. In particular, the width of all profiles parallel to $\mathbf{c}^*$ were found to be sharper than found experimentally. It appears that the density of individuals in the expanded parameter subspace is too low and that the distances between first-generation individuals and the region of the reference solution is too large. Some positive observations are also worth noting, however.

In all tests, the correct signs and the relation $|J_a| < |J_b| < |J_c|$ were found within a few generations even though the parameters J_a, J_b and J_c were larger than the reference values by factors 10 to 300. Similarly, the *k* parameters did fulfil the relation $|k_c| \lesssim |k_a| < |k_b|$, the same as found in the reference solution, but exceeded the reference values by factors of 5 to 40. In all calculations, d_{RS} converged to values between −0.03 and −0.05 within a few generations, *i.e.* close to its reference value. Relatively accurate results of d_{RS} have been found without any problems early on in refinements *C*, *D* and *E*, even though only a single lot was used initially to obtain $I_{\rm calc}$. The number of generations required in refinements *A* and *B* to come close to the reference value of d_{RS} and to obtain the correct qualitative relations between the *k* and *J* parameters was smaller than in *C*, but the time span required was higher, because the calculation of the fitness was based on 20 lots.

For the subsequent tests *F* and *G*, only the master computer and four slots of the Sun SPARC cluster were available, *i.e.* wall clock time doubled compared with the reference refinement. In refinement *F*, the range of starting values for the *k* parameters and for J_a and J_b was reduced by a factor of 100 to $[0 \ldots 1]$ and $[0 \ldots 0.01]$, respectively. The range for J_c was reduced to $[-0.1 \ldots 0]$ and that of d_{RS} to $[-0.005 \ldots -0.001]$. As a consequence, the diversity of the population was reduced by a factor of 10^{10} compared with the reference calculation. The refinement was started using a single lot to calculate fitness. After generation 54, the number of lots was increased to 2. After generation 90, *i.e.* after about 30 h wall clock time, the *J* and d_{RS} parameters were very close to their reference values whereas the *k* parameters were not (Table 2); k_a and k_b had increased somewhat and k_c became slightly negative. In order to test whether the converging evolution of d_{RS} and of the *J* parameters could also be extended to the *k* parameters, refinement was continued with twice the number of lots until it reached generation 128. The final number of lots was 150, the same as used towards the end of the reference refinement. After 312 generations (= 45 days wall clock time!), d_{RS} and the *J* parameters had come somewhat closer to the values of the full refinement but the *k* parameters did not improve significantly. The *R* value dropped to 0.159 (3), only slightly higher than the reference value of 0.148 (3). Inspection of the difference intensities showed that refinement *F* reproduces diffuse features with an accuracy comparable with that of the reference refinement, except for the asymmetries in the experimental profiles along $\mathbf{c}^*$ (Fig. 4*c*). This is a clear indication that the results from test *F* are inferior to the reference solution, *i.e.* calculation *F* became trapped in a local minimum. The observation that a relatively small volume of starting values led to significantly better results than did a relatively large volume should not be generalized, because the evolution of a refinement depends strongly on the shape of the associated χ^2 surface.

Given the improvement in performance and the quality of the results of test calculation *F*, which started with a single lot instead of 20 lots, the influence of the number of lots on the early stages of structure determination was investigated in more detail. The evolutionary process was repeated with the same parameter ranges as in the reference refinement, but now using only a single lot in the beginning (test refinement *G*). The number of lots was doubled at generation 51, 113 and 167 to a total of eight after which no further progress could be observed. Agreement between observed and calculated diffraction features is good. Results after 60 and 209 generations (= 14 and 70 h wall clock time, respectively) are listed in Table 2. The parameter values and their standard deviations correspond roughly to the results of the reference refinement reached after 30 and 50 generations, respectively, *i.e.* after 30 and 54 h of wall clock time (with twice the computing power). This experiment shows that starting DE with a low number of lots helps to approach the region of the global minimum

within a shorter time. In the reference refinement, the number of lots was left constant for large numbers of generations. As a consequence, none or only a few individuals per generation were replaced by children over long sequences of generations. The resulting extended plateau phases show no significant improvement in the fitness of the population. A more automatic change in the number of lots will improve the performance of the refinement considerably. A possible scenario would be to double the number of lots once the number of upgraded individuals falls below a certain level.

5. Conclusions

Determination of disordered crystal structures with a combination of evolutionary algorithms and Monte Carlo simulations comprises several steps. The first and crucial part of the method is the definition of a structural genotype, *i.e.* of a set of parameters describing the disorder adequately. The energy parameters should mimic the atomic interactions in a disordered structure and be minimally correlated. The number should be kept to a manageable size because determining numerical parameters from diffuse scattering is computationally demanding. This task requires intuition, experience and great care. The need to define a model and a considerable computational expense are the main features distinguishing the determination of disordered structures from single-crystal structure analysis. In the latter case, hundreds or thousands of well defined parameters and occupation factors are easily determined in a matter of minutes on a single PC or workstation.

Once an adequate genotype has been defined, the further procedure is relatively straightforward. Differential evolution finds good *approximations* to the disordered structure within a reasonable time, provided that the approximate *range of possible solutions* is known. The number of structures equals the number of individuals in the population. On the one hand, the lack of uniqueness implied in a population may seem a disadvantage. On the other hand, the availability of several solutions with similar fitness increases the probability of finding a structure that fits the diffraction data **and** is sensible from a chemical and physical point of view. This feature may be particularly advantageous when the definition of the genotype happens to be flawed. Welberry and co-workers describe finding approximate structural models by trial and error (Mayo *et al.*, 1999; Welberry, 2000; Welberry *et al.*, 2001). Evolutionary algorithms leave this job to computers and may be expected to identify the region of the global minimum with high probability. In this respect, our method is analogous to modern methods of finding approximate solutions to the phase problem preceding the determination of average crystal structures. Differential evolution is capable of operating with very noisy calculated intensities. This feature allows working with small numbers of lots, accelerates Fourier transformation and facilitates escape from local minima. In contrast, automatic Monte Carlo methods using numerical derivatives with respect to the model parameters depend on a large number of lots from the very beginning to reach a low level of noise in the calculation of χ^2. They would seem less suitable for testing approximate structure solutions.

In principle, refinement of approximate structure models could also be done by differential evolution. However, our calculations show that the performance of the evolutionary process decreases dramatically when the number of lots has to be increased in order to reduce noise in $I_{\rm calc}$. A least-squares technique like automatic Monte Carlo refinement may be more efficient for final optimization of interaction parameters.

The computational effort implied in our implementation of the differential evolution/Monte Carlo technique is considerable. However, the performance of differential evolution can be improved in several respects: the different strategies of differential evolution and the effect of crossover and mutation constants will have to be compared, the number of lots will have to be adapted to the progress of differential evolution in an intelligent way, Fourier transformation of large numbers of lots will have to be parallelized to a much higher degree. Nowadays, clusters of several hundred inexpensive personal computers working in parallel are increasingly common tools in computationally intensive research projects.

It has been shown previously that a complete set of diffuse scattering data of high quality can be collected at a synchrotron in about a day (Weber *et al.*, 2001). At present, the bottleneck is interpreting them. In this study, we hope to have shown that the differential evolution/Monte Carlo technique is a step towards a more efficient and automated determination of disordered crystal structures. We conclude that many more problems in diffuse scattering and disordered structure determination will become practical if computing efficiency can be improved by one order of magnitude.

The authors thank Professor Reinhard Neder, University of Würzburg, for providing computer resources from his laboratory and the team of the Swiss Norwegian Beamlines (ESRF, Grenoble) for help with the experiments.

References

Chang, G. & Lewis, M. (1997). *Acta Cryst.* D**53**, 279–289.
Estermann, M. A. & Steurer, W. (1998). *Phase Transit.* **67**, 165–195.
Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning.* Reading, MA: Addison-Wesley.
Guinier, A. (1963). *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies.* New York: Freeman.
Harris, K. D. M., Johnston, R. L. & Kariuki, B. M. (1998). *Acta Cryst.* A**54**, 632–645.
Kariuki, B. M., Serrano-Gonzàlez, H., Johnston, R. L. & Harris, K. D. M. (1997). *Chem. Phys. Lett.* **280**, 189–195.
Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983). *Science*, **220**(4598), 671–680.
Knorr, K. & Mädler, F. (1999). *J. Appl. Cryst.* **32**, 902–910.
König, O., Bürgi, H. B., Armbruster, T., Hulliger, J. & Weber, T. (1997). *J. Am. Chem. Soc.* **119**, 10632–10640.
Landree, E., Collazo-Davila, C. & Marks, L. D. (1997). *Acta Cryst.* B**53**, 916–922.
Mayo, S. C., Proffen, T., Bown, M. & Welberry, T. R. (1999). *J. Appl. Cryst.* **32**, 464–471.

Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs.* Berlin/Heidelberg/New York: Springer-Verlag.
Neder, R. B. & Proffen, T. (1999). *User Guide DISCUS Version 3.3.* University of Würzburg, Germany.
Price, K. & Storn, R. (1997). *Dr Dobb's J.* (April), pp. 18–24.
Proffen, T. & Neder, R. B. (1997). *J. Appl. Cryst.* **30**, 171–175.
Proffen, T. & Welberry, T. R. (1997). *Acta Cryst.* A**53**, 202–216.
Proffen, T. & Welberry, T. R. (1998). *Phase Transit.* **67**, 373–397.
Rosshirt, E., Frey, F., Boysen, H. & Jagodzinski, H. (1985). *Acta Cryst.* B**41**, 66–76.
Scheidegger, S., Estermann, M. A. & Steurer, W. (2000). *J. Appl. Cryst.* **33**, 35–48.
Shankland, K., David, W. I. F. & Csoka, T. (1997). *Z. Kristallogr.* **212**, 550–552.
Ulyanenkov, A., Omote, K. & Harada, J. (2000). *Physica (Utrecht) B*, **283**, 237–241.
Weber, T., Boysen, H. & Frey, F. (2000). *Acta Cryst.* B**56**, 132–141.
Weber, T., Estermann, M. A. & Bürgi, H. B. (2001). *Acta Cryst.* B**57**, 579–590.
Weber, T., Neder, R. B. & Bürgi, H. B. (2002). In preparation.
Welberry, T. R. (2000). *Acta Cryst.* A**56**, 348–358.
Welberry, T. R., Goossens, D. J., Edwards, A. J. & David, W. I. F. (2001). *Acta Cryst.* A**57**, 101–109.
Welberry, T. R. & Proffen, T. (1998). *J. Appl. Cryst.* **31**, 309–317.
Welberry, T. R., Proffen, T. & Bown, M. (1998). *Acta Cryst.* A**54**, 661–674.
Wormington, M., Panaccione, C., Matney, K. M. & Bowen, K. (1999). *Philos. Trans. R. Soc. London Ser. A*, **357**, 2827–2848.