

## Application Note SC-XRD 503

# Intrinsic Phasing – Advanced Structure Solution in APEX3

- Structure solution for small molecules

### Introduction

The diffraction pattern produced by the interaction of X-rays with a crystal can be understood as the Fourier transform of the complete electron density of the crystal structure. In principle the electron density can be simply reconstructed by Fourier synthesis. In reality, only the intensities of the waves of the diffraction pattern can be recorded.

The information about the phase of the wave, the relative timing when each wave hits, is lost in the

experiment. The problem of elucidating the phases when only the magnitudes are known is referred to as the “phase problem”.

Crystallographers consider a crystal structure to be solved when the phases of enough reflections are sufficiently determined so that most, if not all, of the atoms in the unique part of the unit cell are revealed. There are several ways to solve the crystal structure of a small molecule compound.

## Patterson methods

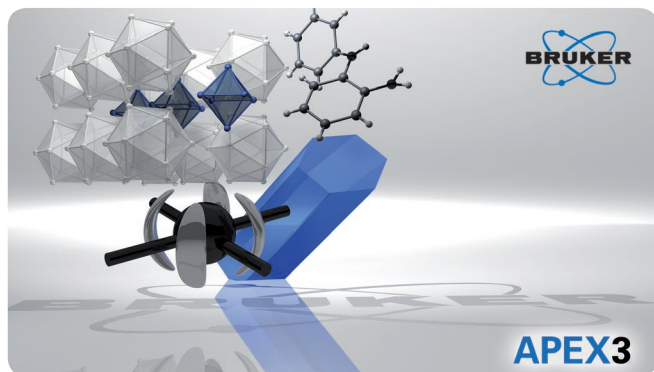
Patterson methods<sup>1,2</sup> employ Fourier maps with the squared amplitudes of the structure factor as coefficients and phase angles all assumed to be 0°. These maps give peaks corresponding to all vectors between any given pair of points. The peak heights are proportional to the atomic numbers of atoms in the structure and the peaks can be used to derive the positions of the atoms. Traditionally, Patterson methods were the first choice for the solution of structures containing a few heavy atoms.

## Direct methods

The term “direct methods” refers to a statistical approach that tries to solve the phase problem by limiting the possible relationships between phase and intensity to a range of probabilities.<sup>3,4</sup> The key point behind the direct method derives from invoking prior structural knowledge, usually that the crystal is composed of discrete atoms, with atoms being small and distinct points relative to the spaces between them.

## Fragment search by integrated Patterson and Direct Methods

This algorithm locates a fragment of known geometry by integrated Patterson, packing, and direct methods<sup>5</sup>. It provides a rotation search that can find the orientation of a search model of any size and allows one torsional degree of freedom. A translation search may locate up to two independent fragments of any size and can take known atoms at fixed positions into account.



The APEX3 software suite is shipped with each D8 QUEST and D8 VENTURE system.

## Dual space methods

Shake-and-bake differs from conventional direct methods by employing a minimal function that is optimized through alternate cycles of reciprocal space phase refinement with density modification and real-space filtering that impose the phase constraints that are implicit in real space.

The shake-and-bake technique has been successfully used to determine the structures of several small proteins.<sup>6</sup>

## Charge flipping

Charge flipping<sup>7,8</sup> is also an iterative algorithm for reconstructing electron densities from diffraction amplitudes. It works on high resolution diffraction data using Fourier recycling. Modification in real space changes the sign of charge density below a threshold, while in reciprocal space the Fourier map is modified without any weighting. Compared to traditional direct methods, it requires much less prior information. Charge flipping needs neither symmetry nor an assumption of atoms. This makes it particularly suitable for structure solution of modulated structures and quasicrystals.

## Intrinsic Phasing

Intrinsic Phasing<sup>9</sup> plays to the strengths of direct methods, which can solve structures best in space group *P1*. With the Laue group known, equivalent intensities are averaged and the data is then expanded to *P1*. A departure from classical direct methods is to start the structure solution process not from random phases but from a Patterson superposition minimum function.

The initial phases provided by the structure solution are then used to determine the appropriate space group, which provides the symmetry information for averaging phases to calculate improved electron densities. Further dual-space recycling is performed to improve the quality of the electron density model. Dual-space recycling employs random omit-maps, where a certain percentage of the peaks are randomly omitted and only the remaining atoms are used to calculate phases. The free-lunch algorithm has also been implemented. This uses density modification to calculate phases for reflections that have not been measured completing the data to a given resolution. Atoms are then assigned to the density's maxima.

This method is shown to work amazingly well and is less demanding on completeness and data quality compared to other direct methods.<sup>6</sup>

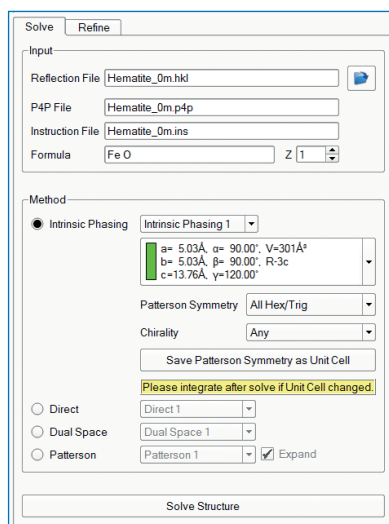


Figure 1: Structure solution module dialog of APEX3 with Intrinsic Phasing solution.

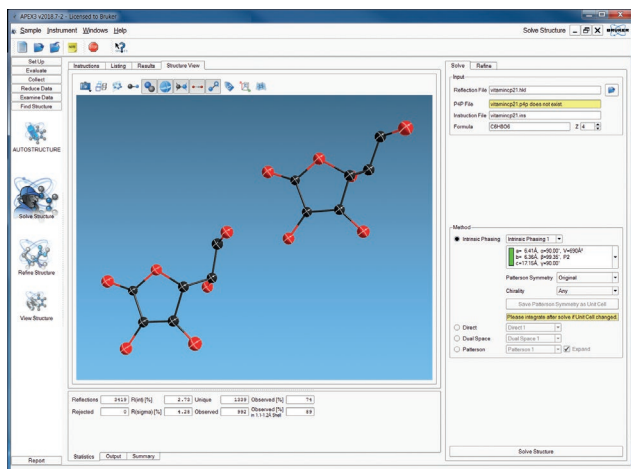


Figure 2: Structure solution module of APEX3: Vitamin C structure displayed as solved with Intrinsic Phasing in the correct space group,  $P2_1$ . All atom types were assigned correctly.

### Structure solution for small molecules with APEX3

APEX3 is a powerful software suite with first class engines for data collection, data integration, structure refinement and reporting. For structure solution and refinement it provides intuitive plug-ins that wrap around George Sheldrick's suite of programs.

Traditional direct methods, Patterson methods, and dual-space methods have been an integral part of the APEX3 Structure Solution module from the beginning, and have provided fast and effective structure solution for years. The latest addition, and the current default method of the solution plugin, is the Intrinsic Phasing method implemented with SHELXT.

In most cases, the default values will give a complete initial model with most of the atom types assigned correctly. The space group search will be limited to the Laue class indicated by the symmetry operators in the instruction file. The space group search can be further limited by choosing between Chiral, Non-centrosymmetric, or any solution. The Intrinsic Phasing solution(s) will be listed in the GUI's combo box towards the bottom of the window, with the

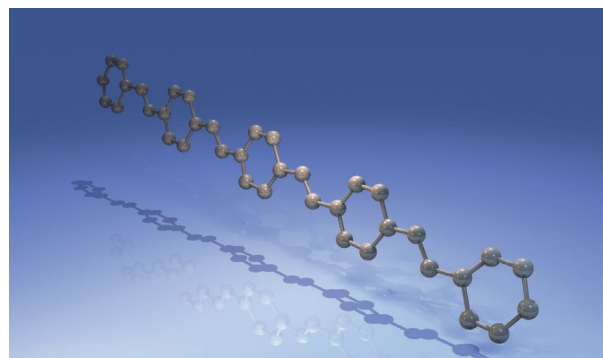


Figure 3: Intrinsic Phasing structure solution in the space group  $P2_1/c$  of polyphenylethylbenzene. The molecule is complete and all atoms were assigned correctly.

best solution on top, together with its reliability criteria  $R1$  and  $\alpha$  (Figure 1). A results file, and an intensity file, the "hkl file", are generated for every solution. Selecting a solution from the combo box will automatically display the structure and results file. On exiting the Structure Solution module, the selected Intrinsic Phasing solution will become the starting input file for the Structure Refinement module.

### Structure solution for small molecules

The following are examples of structures that were solved with the Intrinsic Phasing option of the APEX3 Structure Solution module.

#### Vitamin C – ambiguous space group

Vitamin C crystallizes in the monoclinic space group  $P2_1$ , with two independent molecules in the asymmetric unit. The molecules are related through pseudotranslational symmetry, which systematically lowers the intensities of certain groups of reflections, biasing the E-values statistics toward a value that would be expected for a centrosymmetric space group. This causes the space group determination program to incorrectly choose the centrosymmetric space group  $P2_1/m$ , instead of the correct space group,  $P2_1$ .

Choosing the wrong space group will not deter Intrinsic Phasing from solving the structure correctly in  $P1$ , and subsequently using initial phases to determine the correct space group, presenting the user with the correct solution of the crystal structure in  $P2_1$  (Figure 2).

#### Low resolution, low quality dataset

The structure of a polyphenylethylbenzene compound previously could only be solved with a fragment search approach using PATSEE with a calculated 3D structure of the molecule as the search fragment. Crystals were small, bendable, and film-like, and provided data with high internal reliability criteria to only 1.1 Å. Systematic absences were ambiguous and an instruction file for structure solution was generated with the space group  $P2_1$ . SHELXT was able to solve the structure in  $P2_1/c$  and assign all atoms correctly (Figure 3).

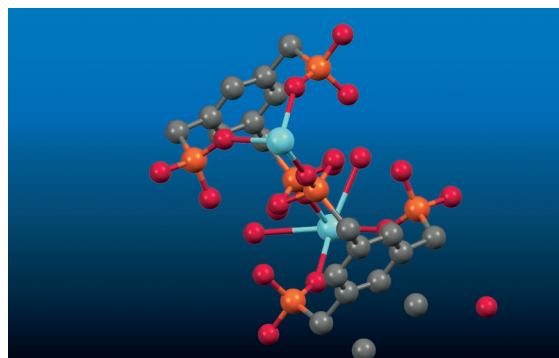


Figure 4: Intrinsic Phasing structure solution in the space group  $P2_1/m$  of  $[La(H_4bmt)(H_5bmt)(H_2O)_2] \cdot 3H_2O$ . All atoms were assigned correctly, except for the disordered lattice water molecules.

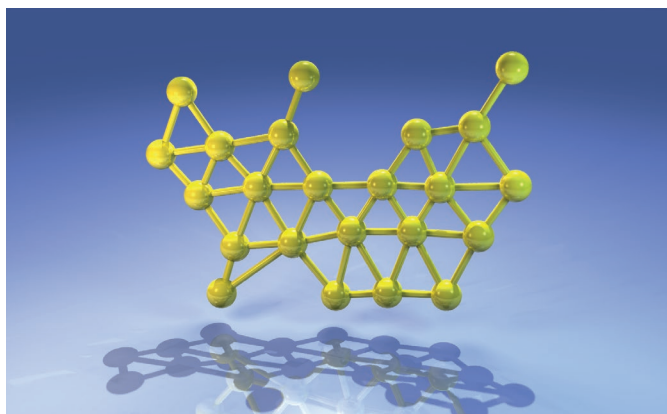


Figure 5: “Chicken-wire” pseudosolution of difluoroxanthone from traditional structure solution methods.

### Weakly diffracting two micrometer-sized MOF Crystal

The structure of a small, 0.002 x 0.013 x 0.020 mm, metal-organic framework crystal,  $[\text{La}(\text{H}_4\text{bmt})(\text{H}_5\text{bmt})(\text{H}_2\text{O})_2]3\text{H}_2\text{O}$ , was solved with Intrinsic Phasing, confirming the unusual space group  $P2_1/m$ , with two formula units ( $\text{C}_{18}\text{H}_{28}\text{LaO}_{23}\text{P}_6$ ) in the unit cell. All atoms were correctly assigned, except for partially occupied lattice water (Figure 4).

### “Chicken-wire” pseudosolutions with flat molecules

Traditional direct and dual space methods provide the dreaded “chicken-wire” pseudosolution that can be observed with certain flat molecules (Figure 5). Where other methods failed to provide the correct solution of difluoroxanthone, Intrinsic Phasing solved the structure correctly in the space group  $P2_1/c$  with all atoms found and correctly assigned (Figure 6).

### Conclusion

APEX3 provides an easy-to-use wrapper around George Sheldrick’s suite of programs for structure solution. The latest addition is SHELXT, implemented as Intrinsic Phasing, with powerful algorithms for structures that were previously difficult to solve, as well as for routine structures. Initial atom assignments are made with a high degree of reliability and the structure is ready for refinement.

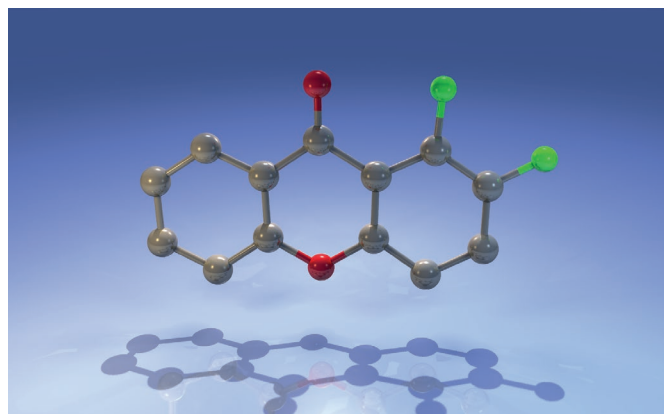


Figure 6: Intrinsic Phasing structure solution in the space group  $P2_1/c$  of difluoroxanthone. The molecule is complete, and all atoms were assigned correctly.

### References

- 1 A. L. Patterson, *Z. Krist.*, **1935**, *A90*, 517-542
- 2 Implemented with: SHELXS (Sheldrick, **1990**)
- 3 J. Karle & H. Hauptman, *Acta Cryst.*, **1950**, *3*, 181
- 4 Implemented with: MULTAN (Main et al., **1980**), SHELXS (Sheldrick, **1990**), SAYTAN (Debaerdemaeker et al., **1985**) and SIR (Burla et al., **1989**)
- 5 Implemented with: PATSEE (Egert, **1985**)
- 6 Implemented with: SnB (Miller et al., **1994**; Weeks & Miller, **1999**), and SHELXD (Sheldrick, **1997, 1998**)
- 7 G. Oszlanyi & A. Suto, *Acta Cryst.* **2004**, *A60*, 134-141 and *Acta Cryst.* **2005**, *A61*, 147-152
- 8 Implemented with: SUPERFLIP (Palatinus, **2004**).
- 9 Implemented with: SHELXT (Sheldrick, **2013**)

### Authors

Dr. Michael Ruf, Bruker AXS, Madison, USA  
Dr. Bruce C. Noll, Bruker AXS, Madison, USA

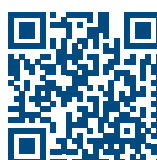
● **Bruker AXS GmbH**

info.baxs@bruker.com

[www.bruker.com](http://www.bruker.com)

**Worldwide offices**

[bruker.com/baxs-offices](http://bruker.com/baxs-offices)



**Online information**

[bruker.com/sc-xrd](http://bruker.com/sc-xrd)

