Solving a Structure in the Reciprocal Space, Real Space and Both by Using the EXPO Software

Angela Altomare *, Nicola Corriero, Corrado Cuocci, Aurelia Falcicchio and Rosanna Rizzi

Institute of Crystallography-CNR, Via G. Amendola 122/o, 70126 Bari, Italy; nicola.corriero@ic.cnr.it (N.C.); corrado.cuocci@ic.cnr.it (C.C.); aurelia.falcicchio@ic.cnr.it (A.F.); rosanna.rizzi@ic.cnr.it (R.R.)

* Correspondence: angela.altomare@ic.cnr.it; Tel.: +39-080-592-9155

Received: 4 December 2019; Accepted: 29 December 2019; Published: 31 December 2019

Abstract: The solution of crystal structures from X-ray powder diffraction data has undergone intense development in the last 25 years. Overlapping, background estimate, and preferred orientation are the main difficulties in the process of determining a crystal structure from the analysis of the mono-dimensional powder diffraction pattern. EXPO is a well-known computer program that is designed to solve organic, inorganic, and metal–organic structures by powder diffraction data, employing the two most widely used kinds of solution methods: Direct Methods, working in the reciprocal space; and Simulated Annealing, working in the real space. EXPO also allows the suitable combination of these two approaches to validate the structural solution. In this paper, we give examples of structure characterization by EXPO with the aim of suggesting a solution strategy for the application of reciprocal-space methods or real-space methods, or both.

Keywords: powder structure solution; EXPO software; reciprocal-space methods; real-space methods

1. Introduction

Solving crystal structures using powder diffraction data is limited with respect to a single crystal because of the presence of diffraction peak overlap, the uncertain estimation of the background and the possible preferred orientation, which can prevent the complete and correct interpretation of the experimental diffraction pattern. Powder solution is extensively applied to small structures, although to much lesser degree to macromolecules. In the last 25 years, there has been growing interest in powder diffraction: innovative theories, methodological approaches, experimental and computational abilities, and advanced software have been developed. Such progress is also reported in the recent publication of the International Tables for Crystallography, Volume H, Powder Diffraction [1], which covers all the aspects of powder diffraction and is the key reference for all researchers involved in this area. The Tables also point out the important role of powder diffraction in crystal structure determination.

In this scenario, a relevant contribution has come also from the software EXPO [2], which is used to solve organic, inorganic, and metal–organic structures. EXPO is a computer program which was designed more than 20 years ago as a single-crystal-like software for the ab-initio solution by powder diffraction data and has been progressively improved by robust theory, methodological approaches, and computational and graphical tools. An important factor is the capacity of the program to carry out all the steps of the powder solution process: the determination of cell parameters and the space group; the extraction of experimental structure factor moduli; solution by Direct Methods and/or Simulated Annealing; refinement by the Rietveld method; and editing and visualizing molecules. One or more of these steps can be executed by following standard processes, which are mainly automatic, or non-default approaches, whose application is easily guided by the EXPO’s user-
friendly graphic tools. The authors of EXPO are also active in the development of computational methods which are able to minimize the execution time of the solution process: the solution by Direct Methods is usually performed in only a few minutes, while the solution by Simulated Annealing can be slower and even require days in the case of quite complex structures. The standard approaches, which are the result of intensive test work on hundreds of structures which have already been solved and published and which are of different structure complexity, suggest solution strategies which are successful in most cases. Several options for non-default procedures are available in the case of the failure of the default run. This paper is focused on applications of EXPO for a solution in the reciprocal space [3]—specifically by Direct Methods (default choice)—or in the real space [4]—in particular by Simulated Annealing—or by both methods.

The decision regarding the best solution strategy mainly depends on the data quality, especially the overall diffracted intensity, experimental resolution and peak overlap; the structure complexity expressed in terms of the number of non-hydrogen atoms in the asymmetric unit and/or degrees of freedom of the structure; and available information on the expected structure model. In particular, peak overlap [5] is unavoidably present and can make the interpretation of a powder diffraction pattern more difficult. Altomare et al. [6] proposed a method for evaluating the overlapping degree in a powder profile, warning that the larger the overlap, the smaller the number of independent observations, with consequences on the rate of success of Direct Methods solution procedures and least-squares refinement processes.

In this paper, we discuss and compare the different conditions requested for a successful solution by Direct Methods and Simulated Annealing, respectively, in EXPO, and we present two case studies that clarify the opportunity of using a specific method rather than the other and explain the strength of using both, if possible. In particular, we show that we can originally and profitably exploit the structure model obtained at the end of the solution process by Simulated Annealing to improve the solution by Direct Methods and confirm the correctness of the Simulated Annealing model.

2. Solving Structures in the Reciprocal Space by EXPO: Case Studies

Solving structures by X-ray powder diffraction data through reciprocal-space methods [3] is a single-crystal-like process. In this perspective, the structure factor moduli are extracted from the experimental diffraction pattern; then, Direct Methods, Patterson Methods, Maximum Entropy Methods or Charge Flipping determine the phases of reflections, finally deriving the electron density map. The structure solution process in the reciprocal space by EXPO consists of a very fast sequence of steps: the extraction, from the powder profile, of the experimental structure factor moduli using the Le Bail method [7]; progressive use of the moduli by Direct Methods to solve the phase problem and statistically and probabilistically evaluate the phases of structure factors; the calculation of the inverse Fourier transform of structure factors as determined in both moduli and phases; the interpretation of the peak positions and intensities in the electron density map to obtain the structure model; and the optimization of the structure model by least-squares-Fourier-recycling methods. In addition to the speed, another great advantage of the Direct Methods ab-initio solution is the minimal need for necessary initial information, as only the chemical formula and diffraction pattern are required. Moreover, when a chemically plausible structure model is obtained by Direct Methods, its reliability is usually high. The solution by Direct Methods is the automatic choice in EXPO because is fast and, if successful, reliable.

The attainment of the correct solution by Direct Methods requires conditions that, unfortunately, are not always fulfilled: i.e., well estimated structure factor moduli extracted from the diffraction pattern and atomic experimental resolution.

Because of the diffraction peak overlap, the accuracy of the extracted moduli is usually poor, even if it can be improved by the use of prior information in the Le Bail algorithm [8] in EXPO. Peak overlap can be reduced, but not completely eliminated. In this paragraph, we consider the set of 155 test structures contained in our database of structures which have already been solved (by using different available methods and/or software) and published, and for which the capacity of solution
by EXPO, by using default and non-default strategies, has been proved. The set is made up of organic, metal–organic, and inorganic structures and covers a broad range of research fields. In Table 1, we provide the experimental resolution (RES), the structural complexity defined by the number of non-hydrogen atoms in the asymmetric unit (NA-noH), and the type of data (laboratory X-ray or synchrotron; see below for the content in the last column). Molecular formula and references corresponding to each test structure are listed as supplementary information in Table S1.

**Table 1.** For each of the 155 test structures, experimental resolution (RES); number of non-hydrogen atoms in the asymmetric unit (NA-noH); type of data (laboratory X-ray or synchrotron data); ratio between the number of independent reflections (IR) and NA-noH (IR/NA-noH).

| Test Structure | RES | NA-noH | Type of Data | IR/NA-noH | Test Structure | RES | NA-noH | Type of Data | IR/NA-noH |
|----------------|-----|--------|--------------|-----------|----------------|-----|--------|--------------|-----------|
| X: Laboratory X-ray Data | S: Synchrotron Data |
| 1 | 2.83 | 64.00 | | 1.00 | 7 | 10.90 | 18.40 | | |
| 2 | 1.43 | 17.45 | | 1.14 | 2 | 24.00 | | | |
| 3 | 1.43 | 18.00 | | 1.14 | 15 | 25.39 | | | |
| 4 | 1.43 | 12.94 | | 1.01 | 17 | 23.43 | | | |
| 5 | 1.53 | 10.90 | | 1.43 | 23 | 20.95 | | | |
| 6 | 1.53 | 10.90 | | 1.43 | 55 | 23.67 | | | |
| 7 | 1.53 | 10.90 | | 1.43 | 6 | 19.13 | | | |
| 8 | 1.43 | 10.90 | | 1.43 | 30 | 19.40 | | | |
| 9 | 1.43 | 10.90 | | 1.43 | 36 | 19.60 | | | |
| 10 | 1.43 | 10.90 | | 1.43 | 36 | 19.80 | | | |
| 11 | 1.43 | 10.90 | | 1.43 | 36 | 20.00 | | | |
| 12 | 1.43 | 10.90 | | 1.43 | 36 | 20.20 | | | |
| 13 | 1.43 | 10.90 | | 1.43 | 36 | 20.40 | | | |
| 14 | 1.43 | 10.90 | | 1.43 | 36 | 20.60 | | | |
| 15 | 1.43 | 10.90 | | 1.43 | 36 | 20.80 | | | |
| 16 | 1.43 | 10.90 | | 1.43 | 36 | 21.00 | | | |
| 17 | 1.43 | 10.90 | | 1.43 | 36 | 21.20 | | | |
| 18 | 1.43 | 10.90 | | 1.43 | 36 | 21.40 | | | |
| 19 | 1.43 | 10.90 | | 1.43 | 36 | 21.60 | | | |
| 20 | 1.43 | 10.90 | | 1.43 | 36 | 21.80 | | | |
| 21 | 1.43 | 10.90 | | 1.43 | 36 | 22.00 | | | |
| 22 | 1.43 | 10.90 | | 1.43 | 36 | 22.20 | | | |
| 23 | 1.43 | 10.90 | | 1.43 | 36 | 22.40 | | | |
| 24 | 1.43 | 10.90 | | 1.43 | 36 | 22.60 | | | |
| 25 | 1.43 | 10.90 | | 1.43 | 36 | 22.80 | | | |
| 26 | 1.43 | 10.90 | | 1.43 | 36 | 23.00 | | | |
| 27 | 1.43 | 10.90 | | 1.43 | 36 | 23.20 | | | |
| 28 | 1.43 | 10.90 | | 1.43 | 36 | 23.40 | | | |
| 29 | 1.43 | 10.90 | | 1.43 | 36 | 23.60 | | | |
| 30 | 1.43 | 10.90 | | 1.43 | 36 | 23.80 | | | |
| 31 | 1.43 | 10.90 | | 1.43 | 36 | 24.00 | | | |
| 32 | 1.43 | 10.90 | | 1.43 | 36 | 24.20 | | | |
| 33 | 1.43 | 10.90 | | 1.43 | 36 | 24.40 | | | |
| 34 | 1.43 | 10.90 | | 1.43 | 36 | 24.60 | | | |
| 35 | 1.43 | 10.90 | | 1.43 | 36 | 24.80 | | | |
| 36 | 1.43 | 10.90 | | 1.43 | 36 | 25.00 | | | |
| 37 | 1.43 | 10.90 | | 1.43 | 36 | 25.20 | | | |
| 38 | 1.43 | 10.90 | | 1.43 | 36 | 25.40 | | | |
| 39 | 1.43 | 10.90 | | 1.43 | 36 | 25.60 | | | |
| 40 | 1.43 | 10.90 | | 1.43 | 36 | 25.80 | | | |
| 41 | 1.43 | 10.90 | | 1.43 | 36 | 26.00 | | | |
| 42 | 1.43 | 10.90 | | 1.43 | 36 | 26.20 | | | |
| 43 | 1.43 | 10.90 | | 1.43 | 36 | 26.40 | | | |
| 44 | 1.43 | 10.90 | | 1.43 | 36 | 26.60 | | | |
| 45 | 1.43 | 10.90 | | 1.43 | 36 | 26.80 | | | |
| 46 | 1.43 | 10.90 | | 1.43 | 36 | 27.00 | | | |
| 47 | 1.43 | 10.90 | | 1.43 | 36 | 27.20 | | | |
| 48 | 1.43 | 10.90 | | 1.43 | 36 | 27.40 | | | |
Figure 1 gives details regarding the errors for the structure factor moduli estimates extracted from the experimental powder profile: for each test structure reported in Table 1, the y-axis displays the corresponding $R_F$ reliability parameter (the structures have been ranked by the increasing value of $R_F$).

$$R_F = \frac{\sum_h |F_h|_{\text{ext}} - |F_h|_{\text{true}}}{\sum_h |F_h|_{\text{true}}} \times 100.0,$$

where the summation is over the number of reflections in the experimental diffraction pattern, $|F_h|_{\text{ext}}$ is the structure factor modulus corresponding to the reflection $h$ extracted by the Le Bail method in EXPO and $|F_h|_{\text{true}}$ is the modulus calculated by the correctly refined model. The $R_F$ values, which are calculated in an a posteriori analysis after the structures have been solved, reveal that the errors range from 21% to 91% and that very large errors may occur: the percentage of structures with $R_F > 40\%$ is 62%, while this value is $35\%$ with $R_F > 50\%$. Such a circumstance clarifies the reasons why the *ab-initio* solution by powder is still a challenge.
One of the causes of the uncertainties regarding the experimental structure factor moduli is undoubtedly peak overlap. It can be useful to quantify the peak overlap extent in a powder pattern. We developed a method that evaluates the peak overlap degree—specifically, the percentage of independent observations in an experimental powder diffraction profile [6]—the larger the overlap, the smaller the number of independent observations, and the smaller the rate of successful solution. The percentage value can be used as one of the predictive indications for the success of the solution process by Direct Methods. In Figure 2, the percentage of independent reflections (IRP) is shown for each of the 155 test structures reported in Table 1: on average, an increasing RF decreases IRP, and the correlation between IRP and RF is −0.75. Only 36% of the reported structures have an IRP larger than 50%.

The average values of RF and IRP for synchrotron data structures are 42% and 52%, respectively, while the corresponding values for the laboratory X-ray data structures are 47% and 44%.

---

**Figure 1.** The RF value corresponding to each of the 155 test structures reported in Table 1. The structures have been ranked by the increasing value of RF.

**Figure 2.** The percentage of independent reflections (IRP) value corresponding to each of the 155 test structures reported in Table 1.
In the last column of Table 1, the ratio between the number of independent reflections, obtained by multiplying the number of reflections in the experimental pattern (NRexp) by IRP, and NA-noH is supplied (IR/NA-noH). Direct Methods usually require that the number of reflections actively used in the phasing process (NRact) is at least seven times the number of non-hydrogen atoms in the asymmetric unit to assure a sufficient number of phase relationships and a good quality electron-density map [6]. NRact is usually much smaller than the number of measured reflections: for the test structures in Table 1, 10% < (NRact/NRexp) × 100 < 46%. This means that when IRP is small, the condition requested by Direct Methods is hardly satisfied.

When we are approaching the solution process of a new structure by Direct Methods, a preliminary estimate of the success probability can be useful. The possibility of success increases with increasing IRP and IR/NA-noH and decreasing RES and NA-noH. According to the analysis of our test structures, Table 2 schematizes the limits—which are empirically determined—of these conditions for the two extreme cases of high and low success rates, respectively. When a structure does not verify the conditions of high and low success rates, the advice is still to attempt the EXPO _ab-initio_ solution process. A structure which is unsolved by Direct Methods can be successfully solved by real-space methods.

| High Success Rate by Direct Methods | Low Success Rate by Direct Methods |
|------------------------------------|------------------------------------|
| IRP > 60% and RES < 1.2Å and NA-noH < 30 and IR/NA-noH > 15 | IRP < 50% and RES > 1.5Å and NA-noH > 35 and IR/NA-noH < 9 |

Organic structures are most resistant to the solution attempts by Direct Methods because the rapid scattering factor decay of light atoms usually limits the experimental resolution, and therefore two case studies of organic structures are given below. They are outside the limits of the conditions of both high and low success rates.

Firstly, we will consider the case of a 5-(5-nitro furan-2-ylmethylene), 3-N-(2-methoxy phenyl), 2-N’-(2-methoxyphenyl) imino thiazolidin-4-one compound [9] with chemical formula C$_{22}$H$_{17}$O$_6$N$_3$S. X-ray diffraction data, shown in Figure 3, were collected at room temperature (293 K) by using an automated Rigaku RINT2500 laboratory diffractometer (50 KV, 200 mA), equipped with an asymmetric Johansson Ge(111) crystal to select the monochromatic CuKa1 radiation ($\lambda$ = 1.54056 Å). The silicon strip Rigaku D/teX Ultra detector was used. The measurement was scanned for diffraction angles (2$\theta$) ranging from 7° to 70° with a step size of 0.02° and a time of 6 s/step. The value of RES is 1.34 Å.

![Figure 3](image-url). Experimental X-ray diffraction pattern of C$_{22}$H$_{17}$O$_6$N$_3$S.

The structure has been automatically and successfully indexed by EXPO; the triclinical cell parameters are $a$ = 11.476, $b$ = 10.912, $c$ = 8.809 Å, $\alpha$ = 103.663°, $\beta$ = 91.490°, $\gamma$ = 84.140°. Assuming $P$-1 to be the space group, NA-noH = 32. The number of reflections in the experimental pattern is 919 and the number of independent reflections (IR) calculated by EXPO according to [6] is 349 (IRP = 38%).
with an IR/NA-noH of 10.9. With these conditions, the \textit{ab-initio} solution process by Direct Methods in EXPO is not promising, but can be attempted. At any rate, the default run did not provide a feasible solution. Therefore, with the aim of facing the peak overlap problem, which was possible responsible for the failure, a non-default attempt was carried out by using the EXPO strategy based on a random approach in the Le Bail extraction step \cite{10}: a special directive was introduced in the input file for EXPO. The input file of the 5-(5-nitro furan-2-ylmethylene), 3-N-(2-methoxy phenyl), 2-N’-(2-methoxyphenyl) imino thiazolidin-4-one compound for EXPO is given in the Supplementary Materials in Table S2. We derived a reasonable structure model, but it is approximate (false atom positions and chemical labels as well as inaccurate bond distances and angles are present) and incomplete. The model was improved by the application of default \cite{11} and non-default structure model optimization strategies \cite{12} ending with the complete and correct solution. The overall execution time for obtaining the correct Direct Methods solution starting from only the chemical formula and experimental diffraction pattern was about 5 min \cite{Intel(R) Core(TM) i7-4510U CPU @ 2.00 GHz 2.60 GHz}. In an a posteriori analysis, after the structure was solved, we checked, by comparison with the true model, that the error for the structure factor moduli extracted by the Le Bail method was about 46%. The hydrogen atoms were geometrically located in the Direct Methods model by using the tool available in EXPO, and then the structure was refined by the Rietveld method \((R_p = 2.034, R_{wp} = 2.777)\). In Figure 4, the observed, calculated and difference profiles as well as the background are shown.

![Figure 4](image)

**Figure 4.** The result of the Rietveld refinement of 5-(5-nitro furan-2-ylmethylene), 3-N-(2-methoxy phenyl), 2-N’-(2-methoxyphenyl) imino thiazolidin-4-one compound by EXPO: the observed profile (blue line), calculated profile (red line), background (green line) and difference profile (purple line) are shown.

The second case corresponds to 2-(5,6-dimethylimidazo[2,1-b]thiazol-3-yl)-1-morpholinoethanone compound with chemical formula \(C_{13}H_{17}N_{3}O_{2}S\) (shorthand code MORPHOLIN), \(Z’ = 2\), which was solved by single-crystal data \cite{13}. X-ray diffraction data, shown in Figure 5, were collected at room temperature (293 K) by using an automated Rigaku RINT2500 laboratory diffractometer (50 KV, 200 mA), equipped with an asymmetric Johansson Ge(111) crystal to select the monochromatic CuKa1 radiation \((\lambda = 1.54056 \text{ Å})\). The silicon strip Rigaku D/teX Ultra detector was used. The measurement was scanned for diffraction angles \((2\theta)\) ranging from 8° to 70° with a step size of 0.02° and a time of 6 s/step. The value of RES was 1.34 Å.
The cell parameters and the space group have been automatically and successfully determined by EXPO: monoclinic with cell parameters: $a = 14.237$ Å, $b = 19.197$ Å, $c = 10.085$ Å, $\beta = 99.082^\circ$; space group $P2_1/c$. NA-noH is 38. The number of reflections in the experimental pattern is 1175 and the number of independent reflections (IR) calculated by EXPO is 331 (IRP = 28%) with IR/NA-noH equal to 8.7. Trials to solve the structure by Direct Methods in EXPO using default and non-default strategies were unsuccessful so we decided to solve the structure in the real space. In an a-posteriori analysis, we verified that the error on the structure factor moduli extracted by the Le Bail method was nearly 47%.

3. Solving Structures in Real Space by EXPO: A Case Study

A structural solution in the real space [4], developed to overcome the limits of solutions in the reciprocal space, is a valid alternative to Direct Methods, especially when the structure is expected to have molecular geometry information available, and the solution in the real space can be successfully carried out without requiring an experimental atomic resolution or the extraction of the structure factor moduli. The structure complexity is described by degrees of freedom (DoFs) corresponding to the external DoFs of the position (3 DoFs) and orientation (3 DoFs) of each structure fragment in the expected model, and corresponding to the internal DoFs of torsion angles. Countless variations of DoFs are randomly generated, and the corresponding cost function (CF) values are monitored. The CF depends on the agreement between the observed and calculated diffraction profiles. Global optimization methods are used to reach the global minimum of CF, which should correspond to the best real-space solution—possibly the correct structure. The efficiency of the real-space solution process depends on the discrepancy between the starting bond distance and bond angle values with respect to the correct parameters and the number of DoFs. The execution time, which is usually longer than the reciprocal-space solution case, mainly depends on the skill of building up the appropriate starting model, on the number of DoFs and on symmetry. In our experience, structures containing only one fragment with numbers of internal DoFs up to 7 are usually easily solved. For more complex structures, it can be effective to use a powerful computer and parallel version software.

The solution in the real space by EXPO searches the global minimum of CF by Simulated Annealing (SA), which has been suitably modified to improve efficiency and reduce the execution time [14,15]. The CF function used in EXPO is $R_{\text{wp}}$ [2].

The MORPHOLIN compound has been solved by SA in EXPO, although the hydrogen atoms have been neglected. Eliminating the H atoms, which do not contribute significantly to X-ray diffraction, effectively decreases the time to evaluate CF for each trial structure. The angular range $8^\circ < 2\theta < 45.30^\circ$ was used (RES = 2.0 Å). The starting expected model of the title compound was assembled using the sketching facilities of ACD/ChemSketch [16], and the geometry was optimized by using MOPAC [17], which was run using EXPO’s graphical interface. Two molecules were
positioned in the unit cell. In Figure 6, the starting model (Figure 6a) obtained by MOPAC and the SA solution (Figure 6b) are shown.

![Figure 6a](image1.png) ![Figure 6b](image2.png)

**Figure 6.** 2-(5,6-dimethylimidazo[2,1-b]thiazol-3-yl)-1-morpholinoethanone (chemical formula C_{13}H_{17}N_{3}O_{2}S (hydrogen atoms are not considered)): (a) the starting model used in the Simulated Annealing (SA) solution process in EXPO; (b) the SA solution.

A total of 18 DoFs (12 external and six torsion angles for the two molecules) were optimized by EXPO during the SA minimization process. No feasible solution was obtained in a default run (20 times). The global optimization algorithm was run, in a non-standard way, 100 times on a Linux workstation. The number of iterations was increased (niter directive set to 2000 in the EXPO input file) to achieve the optimal crystal structure with a reasonable success rate. The overall time spent on the calculation was about 277 h (Intel(R) Xeon(R) CPU E5-2690 @ 2.90 GHz). The structure of the best solution with the lowest profile cost function (Rwp = 7.34) was selected. The root mean square displacement calculated for all non-H atoms after overlaying it upon the structural solution obtained by single-crystal data [13] was 0.0436 Å. The hydrogen atoms were geometrically located by using the tool available in EXPO, and then the structure was refined by the Rietveld method (Rp = 1.625, Rwp = 2.11). In Figure 7, the observed, calculated and difference profiles as well as the background are shown.

![Figure 7](image3.png)

**Figure 7.** The result of the Rietveld refinement of MORPHOLIN by EXPO: observed profile (blue line), calculated profile (red line), background (green line) and difference profile (purple line) are shown.

4. Solving Structures in Both Real Space and Reciprocal Space by EXPO for Solution Validation

After a feasible solution has been attained in the real space, we can go back to the reciprocal space to confirm the SA solution also in the reciprocal space: such an occurrence strengthens the
reliability of the SA model, which can be then moved to the final structure refinement step. In EXPO, the information derived in the real space can be suitably exploited in the reciprocal space by the Le Bail method [8], which is used to extract the structure factor moduli from the experimental diffraction profile. In the absence of additional information—that is, in a standard way—the Le Bail formula iterates starting with arbitrary integrated intensity values and finally estimates the experimental structure factor moduli, the success of which mainly depends on the peak overlap level in the pattern. By default, the Le Bail algorithm tends to equiportion the intensity of strongly overlapping reflections. On the other hand, the relevant property of the Le Bail method is that it is sensitive to the starting point: it can improve the results when additional information, in some way close to the true structure solution, is used to start the Le Bail iteration [8]. We can use the SA solution as a priori information: instead of starting with arbitrary integrated intensity values, the intensities calculated by the SA model are used to activate the Le Bail extraction process. If the calculated intensities are close to the true values, the error of the extracted moduli is reduced and the probability of solving the structure by Direct Methods is increased.

We adopted this approach with MORPHOLIN and actively used the SA model shown in Figure 6b to derive the starting integrated intensity values for use in the Le Bail iterations (this can be easily carried out by EXPO by selecting a few graphical steps). The newly extracted structure factor moduli were used in a successive solution step by Direct Methods, which provided a feasible solution in about 2 min (Intel(R) Core(TM) i7-4510U CPU @ 2.00 GHz 2.60 GHz). The root mean square displacement calculated for all non-H atoms overlaying Direct Methods and SA models was 0.041 Å. It is worth noting that such a result confirms that the SA solution is correct, as it is able to improve the Le Bail extraction. Indeed, in an a posteriori analysis, we verified that recycling the SA information in the Le Bail method reduced the error of the extracted moduli from 47% to 36%, enabling the correct solution.

The confirmation by Direct Methods is an effective step for validating the solution obtained by Simulated Annealing and allows the progression to the final step of the Rietveld refinement.

5. Conclusion

In the last 25 years, remarkable progress has been achieved in the structure solution process by using X-ray powder diffraction data, but the range of success must still be improved. We can decide to adopt methods which work in the reciprocal space or in the real space, depending on the limits and performances of the methods related to the available information, the characteristics and the level of complexity of the structure under study.

A great advantage offered by the software EXPO is its capacity of solving structures in both the reciprocal space and the real space; in particular, by Direct Methods and/or Simulated Annealing. Firstly, the ab-initio solution by Direct Methods, which requires minimal information, was tested. Several structural model optimization strategies can be applied in EXPO, some of them automatically, some by the user’s request, and all are well supported by the user-friendly graphic interface. They usually succeed in providing the correct structure solution, even starting from a very approximate Direct Methods model. When the reciprocal space solution does not succeed, if the expected geometry of the structure is known, one could try to solve the structure by Simulated Annealing. A feasible successful solution by Simulated Annealing can be exploited to reduce the errors affecting the Direct Methods process and also solve structures in the reciprocal space. A positive outcome from Direct Methods is an undoubted achievement with regards to the success of the solution derived by Simulated Annealing.

Supplementary Materials: The following are available online at www.mdpi.com/2073-4352/10/1/16/s1, Table S1: For each test structure: molecular formula and reference, Table S2: The input file of 5-(3-nitro furan-2-ylmethylen), 3-N-(2-methoxy phenyl), 2-N’-(2-methoxypyphenyl) imino thiazolidin-4-one compound for EXPO.

Author Contributions: A.A. conceived and designed the main ideas, supervised the whole work and wrote a major part of the text; N.C. and A.F. contributed to the applications section; C.C. designed and realized the
Acknowledgments: The research work has been supported by the National Research Council of Italy (CNR).

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References
1. International Tables for Crystallography, Volume H, Powder Diffraction; Gilmore, C.J., Kaduk, J.A., Schenk, H., Eds.; Wiley: New York, NY, USA, 2019.
2. Altomare, A.; Cuocci, C.; Giacovazzo, C.; Moliterni, A.; Rizzi, R.; Corriero, N.; Falcicchio, A. EXPO2013: A kit of tools for phasing crystal structures from powder data. J. Appl. Cryst. 2013, 46, 1231–1235, doi:10.1107/S0021889813013113.
3. Altomare, A.; Cuocci, C.; Moliterni, A.; Rizzi, R. Solving crystal structures using reciprocal-space methods. In International Tables for Crystallography, Volume H, Powder Diffraction; Gilmore, C.J., Kaduk, J.A., Schenk, H., Eds.; Wiley: New York, NY, USA, 2019; pp. 395–413, doi:10.1107/97809553602060000957.
4. David, W.I.F. Real-space methods for structure solution from powder-diffraction data: Application to molecular structures. In International Tables for Crystallography, Volume H, Powder Diffraction; Gilmore, C.J., Kaduk, J.A., Schenk, H., Eds.; Wiley: New York, NY, USA, 2019; pp. 414–432, doi:10.1107/97809553602060000958.
5. Dinnebier, R.E.; Billinge, S.J.L. Overview and principles of powder diffraction. In International Tables for Crystallography, Volume H, Powder Diffraction; Gilmore, C.J., Kaduk, J.A., Schenk, H., Eds.; Wiley: New York, NY, USA, 2019, pp. 2–23, doi:10.1107/97809553602060000935.
6. Altomare, A.; Cascarano, G.; Giacovazzo, C.; Guagliardi, A.; Moliterni, A.G.G.; Burla, M.C.; Polidori, G. On the Number of Statistically Independent Observations in a Powder Diffraction Pattern. J. Appl. Cryst. 1995, 28, 738–744, doi:10.1107/S0021889895009757.
7. Le Bail, A.; Duroy, H.; Fourquet, J.L. Ab-initio structure determination of LiSbWO6 by X-ray powder diffraction. Math. Res. Bull. 1988, 23, 447–452, doi:10.1016/0025-5408(88)90019-0.
8. Altomare, A.; Carrozzini, B.; Giacovazzo, C.; Guagliardi, A.; Moliterni, A.G.G.; Rizzi, R. Solving Crystal Structures from Powder Data. I. The Role of the Prior Information in the Two-Stage Method. J. Appl. Cryst. 1996, 29, 667–673, doi:10.1107/97809553602060007467.
9. Rahmani, R.; Djafrì, A.; Chouaih, A.; Djafrì, A.; Hamzaoui, F.; Rizzi, R.; Altomare, A. Synthesis, molecular and solid state structure of 5-(5-nitro furen-2- ylmethylene), 3-N-(2-methoxy phenyl), 2-N'-(2-methoxyphenyl) imino thiazolidin-4-one: X-ray powder diffraction and DFT studies. J. Mol. Struct. 2017, 1143, 259–264, doi:10.1016/j.molstruc.2017.04.091.
10. Altomare, A.; Giacovazzo, C.; Moliterni, A.G.G.; Rizzi, R. A random procedure for the decomposition of a powder pattern in EXPO. J. Appl. Cryst. 2001, 34, 704–709, doi:10.1107/S0021888901012729.
11. Altomare, A.; Cuocci, C.; Giacovazzo, C.; Moliterni, A.; Rizzi, R. Correcting resolution bias in electron density maps of organic molecules derived by direct methods from powder data. J. Appl. Cryst. 2008, 41, 592–599, doi:10.1107/S0021888908011527.
12. Altomare, A.; Cuocci, C.; Giacovazzo, C.; Moliterni, A.; Rizzi, R. COVMAP: A new algorithm for structure model optimization in the EXPO package. J. Appl. Cryst. 2012, 45, 789–797, doi:10.1107/S002188891201953X.
13. Veltri, L.; Mancuso, R.; Altomare, A.; Gabriele, B. Divergent Multicomponent Tandem Palladium-Catalyzed Aminocarbonylation-Cyclization Approaches to Functionalized Imidazothiazinones and Imidazothioazoles. ChemCatChem 2015, 7, 2206–2213, doi:10.1002/cctc.201500213.
14. Altomare, A.; Corriero, N.; Cuocci, C.; Moliterni, A.; Rizzi, R. The Hybrid Big Bang-Bing Crunch method for solving crystal structure from powder diffraction data. J. Appl. Cryst. 2013, 46, 779–787, doi:10.1107/S0021888913010571.
15. Altomare, A.; Corriero, N.; Cuocci, C.; Falcicchio, A.; Moliterni, A.; Rizzi, R. Direct space solution in the EXPO package: The combination of the HBB-BC algorithm with GRASP. J. Appl. Cryst. 2018, 51, 505–513, doi:10.1107/S1600576718002984.
16. ACD/ChemSketch. Available online: http://www.acdlabs.com/resources/freeware/chemsketch/ (accessed on 31 December 2019).
17. Stewart, J.J.P. MOPAC2016. Stewart Computational Chemistry. Colorado, CO, USA, Springs. 2016. Available online: http://OpenMOPAC.net (accessed on 31 December 2019).

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).