

Frontiers between crystal structure prediction and determination by powder diffractometry

Armel Le Bail

Laboratoire des Oxydes et Fluorures, CNRS UMR 6010, Université du Maine, avenue O. Messiaen, 72085 Le Mans Cedex 9, France

(Received 10 May 2007; accepted 28 February 2008)

The fuzzy frontiers between structure determination by powder diffractometry and crystal structure prediction are discussed. The application of a search-match program combined with a database of more than 60 000 predicted powder diffraction patterns is demonstrated. Immediate structure solution (before indexing) is shown to be possible by this method if the discrepancies between the predicted crystal structure cell parameters and the actual ones are $<1\%$. Incomplete chemistry of the hypothetical models (missing interstitial cations, water molecules, etc.) is not necessarily a barrier to a successful identification (in spite of inducing large intensity errors), provided the search-match is made with chemical restrictions on the elements present in both the virtual and experimental compounds. © 2008 International Centre for Diffraction Data. [DOI: 10.1154/1.2903488]

Key words: crystal structure prediction, powder diffraction, search-match

I. INTRODUCTION

Why does crystal structure prediction (CSP) appear so closely related to powder diffraction methodologies? Should we apply CSP only to compounds already synthesized but existing exclusively as fine powders with unindexed diffraction patterns? Or should we try to enumerate all the most probable crystal structures (organics, inorganics, hybrids) right now? The fact is that structure determination from single crystal data is currently almost totally efficient at solving any problem [proteins with up to 6000 independent non-hydrogen atoms (Burla *et al.*, 2006)], provided one suitable single crystal is available. The consequence is that most efforts in structure solution are now concentrated on cases still unsolved, because of the absence of this essential, large-enough, and well-organized single crystal. A too-complex structure can preclude structure determination by powder diffractometry (SDPD) in spite of a successful indexing (at least for inorganics), since the direct-space methods [recent review by Černý and Favre-Nicolin (2007)] permit relatively easy solution of molecular compounds, provided the molecular formula is well established, whatever its size. However, there are cases where indexing the powder pattern [for a review, see Bergmann *et al.* (2004)] is even not realized at all, so that CSP has a clear and unique role to play. On the other hand, even if showing strong isotropic or anisotropic line broadening (precluding indexing), the experimental powder diffraction pattern remains a powerful fingerprint which can be used for the selection of the adequate model (if any) among the generally too-numerous predictions. We may now feel that CSP is useless when structure determination is possible, and is useful only for otherwise intractable problems. This is a short-term vision because of the current inefficiency of systematic total predictions in arbitrary thermodynamic conditions, a consequence of a poor theory of materials, or at least a poor use of quantum mechanics and chemistry (*ab initio* calculations).

Massive inorganic crystal structure predictions have been done recently. For millions of virtual zeolites (Foster and Treacy, 2003) or related materials, the composition and connectivity are imposed: corner-sharing tetrahedra in SiO_2 ,

(Al/P) O_4 , etc.; this is a search for the more realistic polymorphs. For some other predicted inorganic compounds the exact composition was not strictly imposed, but at least elements were selected and some geometrical rules for organizing them were applied, for instance, by limitation to exclusive corner sharing of polyhedra within the GRINSP software (Le Bail, 2005). This is the very beginning of a promising field, justifying the creation of a database, the PCOD (Predicted Crystallography Open Database), containing more than 60 000 (and rising) crystal data (CIFs available at <http://www.crystallography.net>) of hypothetical silicates, sulfates, phosphates, vanadates, niobates, fluoroaluminates, etc. Predicted crystal structures need to be confirmed, and this has to be done before the structure determination, if one expects any benefit from the efforts expended. The next step was thus to calculate the corresponding powder patterns and to test the feasibility of search-match identification procedures [latest round robin results in Le Meins *et al.* (2003)], allowing for some kind of “immediate structure solution,” even before indexing, which is discussed here.

II. SHORT OVERVIEW OF CSP VERSUS SDPD

Full prediction is supposed to anticipate accurately all the possible crystal structures (chemical content, cell, space group, and atomic positions) before synthesis or discovery in nature. Approaches can be very basic up to highly sophisticated (*ab initio*)—the latter option being still computationally intensive. Since the question “are crystal structures predictable?” by Gavezzotti (1994), or the early reviews of computational studies in solid state chemistry (Catlow, 1997; Catlow *et al.*, 1993), efforts in organic chemistry to predict the packing of previously known molecules (providing polymorphic candidate structures ranked by energy) are progressing slowly through blind tests organized by the CCDC (Motherwell *et al.*, 2002; Day *et al.*, 2005). However, although the use of a previously established molecular formula is invaluable for the search of the structures of possible polymorphs, this approach is unable to predict any new molecule.

There are different levels of prediction depending on the previous knowledge of the system studied. The more we know, the more complex structures we can predict (or solve). We can imagine software combining random elements from the periodic table and producing full predictions in any space group. That dream is perhaps not too far from us now, but the current attempts are much more limited.

A. Predictions knowing the cell and the constituent atoms

This is level zero for CSP; i.e., prediction appears unnecessary here since a structure determination should succeed, either from single crystal or powder diffraction data. Realizing a global optimization though selecting a cost function not using the diffraction data from which the cell was obtained is possible, anyway, and sometimes is the only way (for instance, if the cell is deduced only from the electron diffraction data of an isolated microcrystal in a complex mixture). A genetic algorithm was recently implemented (Woodley, 2004) in the GULP software (Gale, 1997) in order to generate crystal framework structures from the knowledge of the unit cell dimensions and constituent atoms only. The structures of the better candidates are relaxed by minimizing the lattice energy, which is based on the Born model of a solid. Whether this approach can solve structures more easily than the current structure determination methodologies (using diffracted intensities when available) remains to be demonstrated.

B. Predictions knowing the molecule, not the cell

In this case one should speak of molecular packing prediction or molecular polymorph prediction instead of the more general CSP, as claimed in the titles of the CCDC blind tests (Motherwell *et al.*, 2002; Day *et al.*, 2005), which are a bit too general and confusing. Indeed, the molecular packing process involved is able to produce hypothetical polymorphs ranked by energy considerations. The possible three-dimensional packings decide the cells. Some recent examples of successful results are those of pigments (Schmidt *et al.*, 2005; Schmidt *et al.*, 2006; Panina *et al.*, 2007). Generally, not all space groups are explored, limiting the efforts to the less than 20 most frequent space groups observed for organic compounds. Though there is no need for a powder pattern for the building of the models, the confirmation in a real case (since no single crystal is available in principle) has to be performed by comparison with a powder pattern (*a priori* unindexed; otherwise, the list of predicted models could be shortened by comparison with the known cell parameters, and perhaps the direct-space methods would have solved the crystal structure much more easily). Moreover, the question of the selection of the best structure candidate by comparison with powder diffraction data (because the predicted cell parameters may substantially differ from the actual ones) is leading to the establishment of some new similarity indices (Hofmann and Kuleshova, 2005) more efficient than the Rietveld (1969) profile reliability factors, which may fail here. Molecular packing prediction was extended to hybrid materials by the AASBU (Automated Assembly of Secondary Building Units) approach (Mellot-Draznieks *et al.*, 2000, 2002), using Cerius² (Molecular Simulations,

2000) and GULP in a sequence of simulated annealing plus minimization steps for the aggregation of large structural motifs, leading for instance to the solution of two quite complex structures (Férey *et al.*, 2004, 2005). Both have cubic cells with large volumes (380 000 and 702 000 Å³, space group *Fd-3m*, with 68 and 74 independent non-hydrogen atoms). The correct hypothesis (among others) of the presence of super-tetrahedra allowed producing super-zeolites as structure candidates by molecular packing prediction. Not 1 000 000 super-zeolites were obtained, only three, and the correct model was among them (a MTN analog). This can be considered as a quite lucky successful attempt at molecular packing prediction. If the synchrotron patterns had been indexed a bit sooner (something reputed to be easy particularly for even large cubic compounds), then the direct space methods would have certainly also solved the structures by rotations and translations of these super-tetrahedra, which were guessed after considerable efforts (Mellot-Draznieks and Férey, 2005) in constructing various possible building units by “reticular synthesis” (Yaghi *et al.*, 2003).

This raises the question of the number of predictable structures. If more than 1 000 000 zeolites are predicted when less than 20 zeotypes are actually known, what could be the total number of predicted inorganic compounds? We certainly need better ways to select only the correct structure candidates, otherwise we will be submerged soon by billions of structures. There are ~100 000 inorganics in the ICSD, which would expand to 1 000 000 by adding “only” ten new hypothetical polymorphs per known phase, on average, not speaking of completely new formulations. Moreover, molecular structures, on their side, will possibly provide billions of models (just think of the more than 30 000 000 molecules of the Chemical Abstracts Service Registry), so that predicted hybrids will be even much more. This is why solving complex hybrid structures by molecular packing prediction looks extremely lucky. Will such exploits become routine? There are three possible answers to that question, depending on the level of previous knowledge:

1. Yes. If the molecule, the cell, and the space group are known, then direct-space methods need only 50 or 100 reflections from a powder pattern for solving the structure, whatever the cell volume (six degrees of freedom per molecule rotated and translated). But this is not CSP, this is SDPD.
2. Perhaps, by partial prediction (without the cell but with known content). This is the “molecular packing prediction.” Exact cell content knowledge is needed for success, because no Fourier difference synthesis would reveal anything clear from bad quality (since unindexed at the beginning) powder data alone.
3. No. Without cell, content, or any clue, the full total prediction at such complexity level looks totally impossible or would require genius chemists.

C. Predictions based on geometrical rules

Data mining has provided clear topological rules for a long time. If these rules can be described mathematically, then the predictions can be systematic, allowing for a complete enumeration of the possibilities. The most well-known example concerns zeolites: the topology is totally defined by

considering tetrahedra connected exclusively by corners (4-connected 3D net). This is polymorph prediction/ enumeration as well: the exact chemical content is known, and all models have the same SiO_2 [or $(\text{Al}/\text{P})\text{O}_4$, etc.] formula. An exhaustive list of uninodal models was provided and structurally evaluated (Foster *et al.*, 2003), but the mathematical expressions are not established for polynodal zeolites. Without mathematical expression, the exploration can be done by a Monte Carlo approach (Treacy *et al.*, 2004). The structures gathered in the database of $>1\,000\,000$ hypothetical zeolites (Foster and Treacy, 2003) were produced by a 64-bit processor computer cluster grinding away non-stop, generating graphs and annealing them, and the selected frameworks being then re-optimized with the GULP program using atomic potentials. Extending this approach, the GRINSP software (Le Bail, 2005) explored more generally the N -connected 3D nets for $N=3, 4, 5, 6$, and mixed $N-N'$ values. Though the combination of MX_n and $\text{M}'\text{X}_{n'}$ polyhedra is a user choice, the chemical proportions are not imposed there, so that the chemical content is part of the prediction as well (in fact, the topology defines also a restrained list of possible formulations). Another package, SPUDS, is dedicated especially to the detailed prediction (tilting angles, cell, and space group) of perovskites, being one special arrangement among the 6-connected 3D nets (Lufaso and Woodward, 2001).

D. *Ab initio* crystal structure prediction—energy landscape

First-principles studies, the use of parameter-free quantum mechanical calculations in CSP, are too demanding computationally. Relatively simple cases have been investigated, including carbon, CsI, TiO_2 , CsHSO_3 , HCN, and CsGeCl_3 , as undertaken by Milman and Winkler (1999) and the high-pressure structure of LaF_3 (Winkler *et al.*, 2003). This was done with the CASTEP software, using the density functional theory (DFT), applying a pseudo-potential plane-wave code (Payne *et al.*, 1992). In the case of medium to large structures, it is possible to check by the DFT approach the credibility of a few structure candidates. CSP by merging data mining with quantum mechanics was applied to alloy systems (Fischer *et al.*, 2006) in a search for the feasibility of isostructural compounds, not to the prediction of new arrangements. Indeed, the *ab initio* techniques are not yet building candidate structures as fast as the molecular packing polymorph or inorganic structure generators can do. As seen above, checking for some certainty is even more generally done by applying empirical potentials. A concept of “energy landscape” of chemical systems was used for structure prediction with the computer program G42, producing AB_2 compounds (Wevers *et al.*, 1998) for instance, and many others (Schön and Jansen, 2001a, 2001b).

III. DATABASES OF PREDICTED STRUCTURES

The list of software in the above paragraph is small because structure and properties prediction is obviously an unavoidable part of our future in crystallography and chemistry. From these approaches, if the zeolites [inserted into the hypothetical zeolites database (Foster and Treacy, 2003)] and

TABLE I. Current numbers of entries in the PCOD and the P2D2-1, sorted with the chemical formula or combinations of $\text{MX}_n/\text{M}'\text{X}_{n'}$ polyhedra [isostructural series in parentheses] for a total of 16 058 different structure-types, and more than 60 000 hypothetical phases (including isostructural compounds).

Entry numbers	Formula or polyhedra	Isostructural series added
4786	SiO_2	$[(\text{Al}/\text{P})\text{O}_4, (\text{Al}/\text{Si})\text{O}_4 \text{ and } (\text{Al}/\text{S})\text{O}_4]$
4138	AlO_6/BO_3	
2394	VO_5/PO_4	$[\text{VO}_5/\text{SiO}_4, \text{VO}_5/\text{SO}_4, \text{TiO}_5/\text{SiO}_4]$
1747	$\text{TiO}_6/\text{SiO}_4$	$[\text{TiO}_6/\text{PO}_4, \text{TiO}_6/\text{SO}_4, \text{ and substituting Ti by Ga, Nb, V, Zr}]$
1328	TiO_6/VO_5	$[\text{VO}_6/\text{VO}_5]$
1318	V_2O_5	
232	SiO_4/BO_3	
45	B_2O_3	
33	AlF_3	$[\text{FeF}_3, \text{GaF}_3 \text{ and } \text{CrF}_3]$
24	$\text{AlF}_6/\text{CaF}_6$	
13	$\text{AlF}_6/\text{NaF}_6$	

the GRINSP productions [inserted into the PCOD (Le Bail, 2003)] are excluded, the number of inorganic crystal structure predictions is of the order of a few dozens of compounds, some of them being included in the ICSD. On the other hand, organic polymorphs or hybrids predicted by the molecular packing techniques are not yet produced in a systematic way, but only for the characterization of an already existing compound (unindexed powder pattern); the corresponding crystal structures, once solved, are refined from diffraction data and are inserted into the Cambridge Structural Database (CSD) (Allen, 2002). Unless one wants many different researchers performing again and again the same predictions, it is clear that the most probable hypothetical structures should be gathered into databases, and because their number will be quite large, they must be separated from the actual structures.

As the next obvious step, the powder patterns calculated from the PCOD were included into the P2D2-1 (Predicted Powder Diffraction Database version 1) so as to enable identification by search-matching procedures. The P2D2-1 contains more than 60 000 silicates, phosphates, sulfates of Al, Ti, V, Ga, Nb, Zr, or zeolites, fluorides, etc. (Table I). Its utility for automated search-match identification, after coupling with the EVA software (Bruker) is shortly demonstrated here. The P2D2-1 text file contains chemical formula, cell parameters, probable space group, d - I pairs, Miller indices, and $I/I(\text{cor})$ calculated directly from the PCOD CIFs by the CIF2POW software. This text-file was compiled in binary files for compatibility with the EVA search-match program (Bruker).

IV. SEARCH-MATCH THROUGH THE P2D2-1

The searches were performed by using simultaneously the PDF-2 (ICDD) as the “master database” and the P2D2-1 as a “user database.” Two of the four examples below correspond to structures solved by SDPD. Not all predictions will correspond to complete chemical formulations; the frameworks built by corner-sharing polyhedra may contain empty

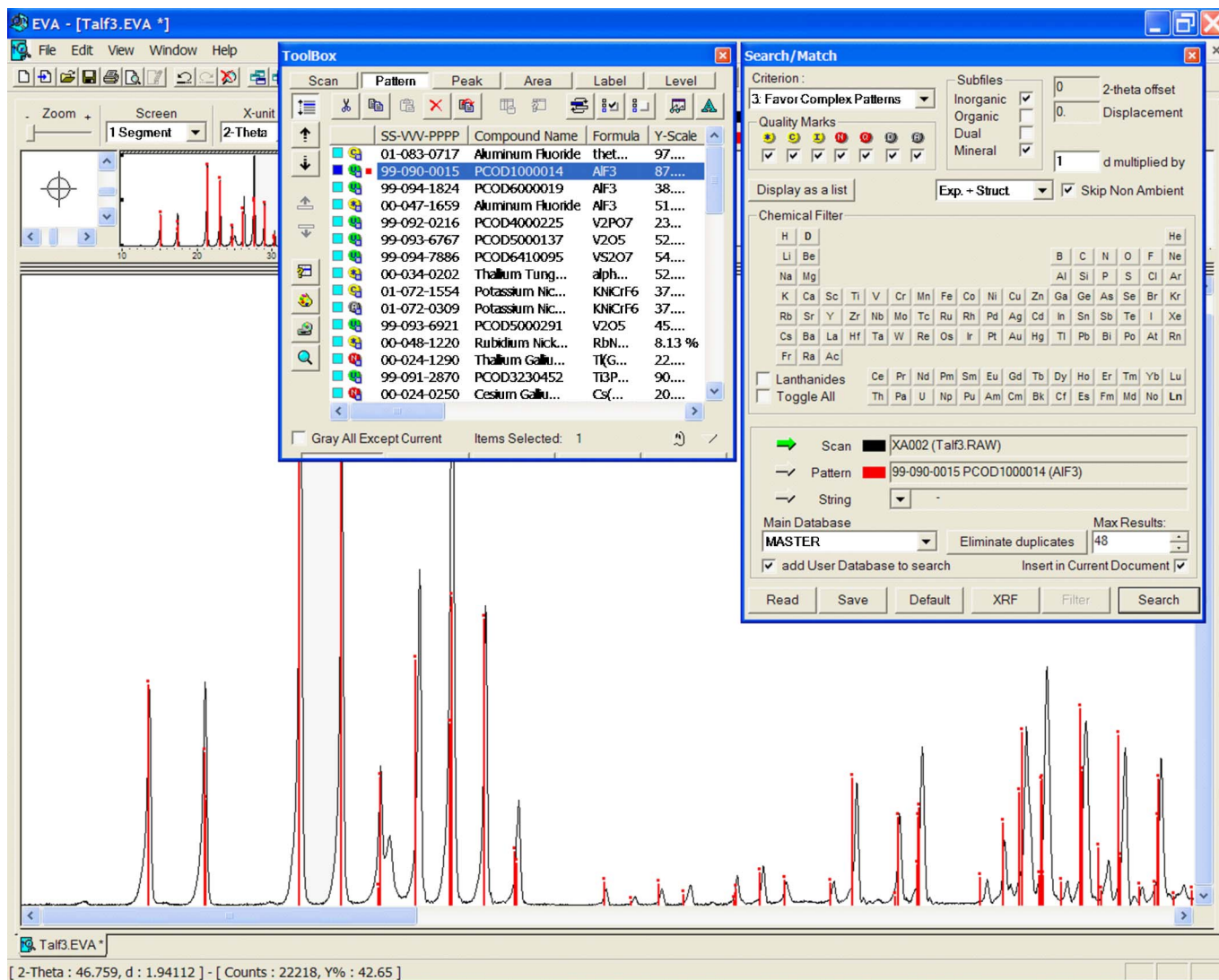


Figure 1. (Color online) Successful search-match attempt by using the EVA (Bruker) software combined with the PDF-2 (Master database) and the P2D2-1 (user database) in the case of τ -AlF₃.

holes or tunnels, and the GRINSP predictions may not even be electrically neutral.

A. Example 1: τ -AlF₃

The GRINSP predictions include 6-connected 3D nets so that one of the latest discovered aluminum fluorides, i.e., τ -AlF₃, the structure of which was determined from powder diffraction data (Le Bail *et al.*, 1992), is among the PCOD/P2D2-1 entries. This is a case where the actual and virtual structures have the same chemical formula, with PAD = 0.52 (percentage of absolute difference on cell parameters, averaged). The actual cell is tetragonal with $a = 10.184$ Å and $c = 7.174$ Å, and the predicted lattice constants being 10.216 Å and 7.241 Å, respectively. A global search with EVA (no chemical restraint) results in the actual compound (PDF-2) in first position and the virtual one (P2D2-1) in second (Figure 1). If the P2D2-1 had existed in 1992, then this new and surprising aluminum fluoride crystal structure would have been directly solved at the identification stage.

B. Example 2: K₂TiSi₃O₉·H₂O

For this second case, a predicted model showing incomplete chemistry when compared to the known compound is chosen, with PAD=0.63. The actual compound is K₂TiSi₃O₉·H₂O, with the structure solved from powder diffraction data (Dadachov and Le Bail, 1997), orthorhombic, $a = 7.136$ Å, $b = 9.908$ Å, and $c = 12.941$ Å. The predicted framework is [TiSi₃O₉], $a = 7.22$ Å, $b = 9.97$ Å, and $c = 12.93$ Å. Without chemical restraint, the correct PDF-2 entry comes at the head of the list, but no virtual model is identified. By using the framework chemical restraint (Ti/Si/O), the correct P2D2-1 entry comes in third position in spite of large intensity disagreements with the experimental powder pattern, since K and H₂O are lacking in the GRINSP-PCOD model (Figure 2). Trying to perform identification from the P2D2-1, it is recommended to test with various angular ranges; the best results may come by limiting the maximum angle to 40 or even 30 ° 2 θ , where peak position discrepancies are less dramatic than at larger angles.

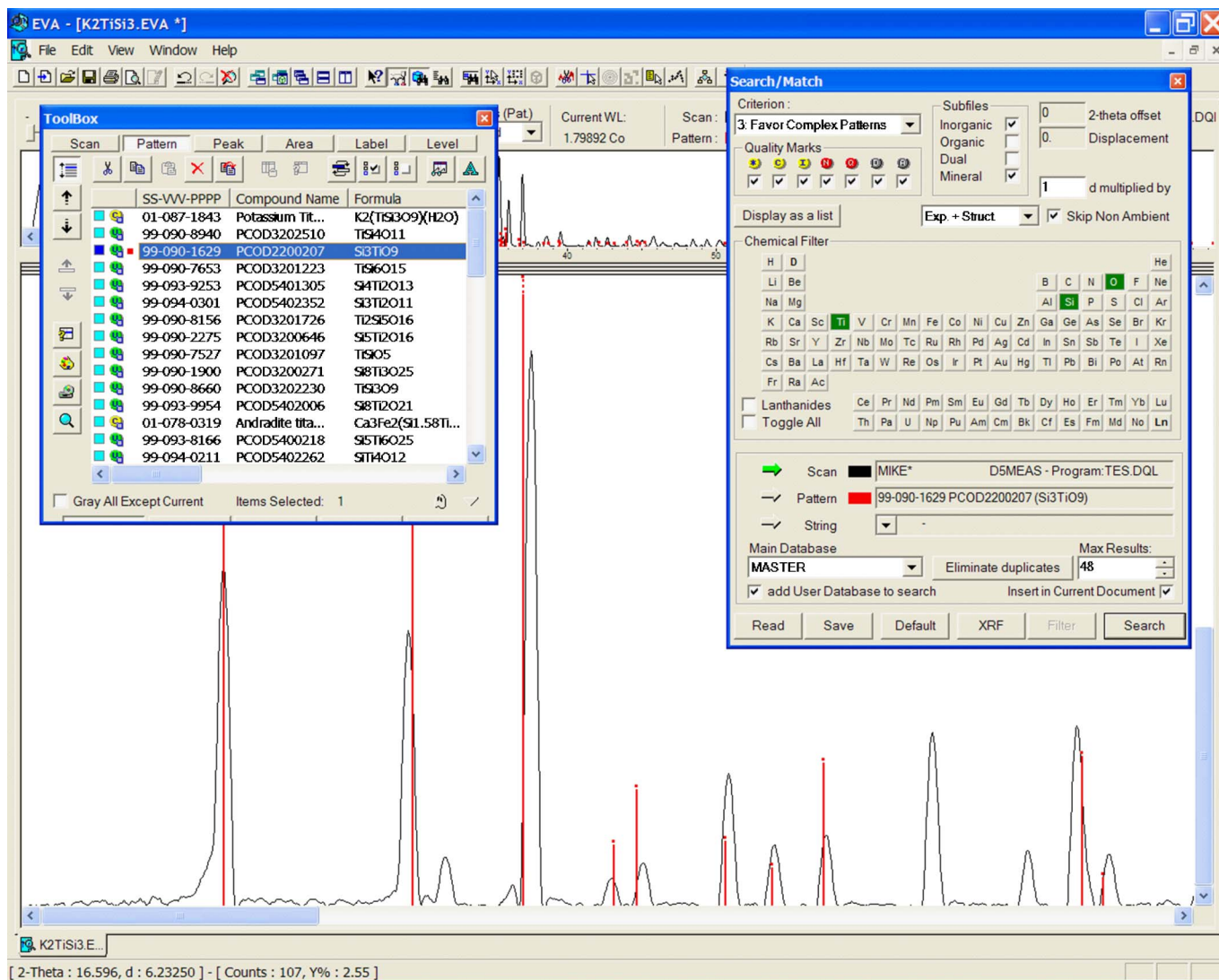


Figure 2. (Color online) Identification of $K_2TiSi_3O_9 \cdot H_2O$ by chemically restraining the search to compounds containing Ti/Si/O in the PDF-2 and the P2D2-1.

C. Example 3: $Na_4Ca_4Al_7F_{33}$

This is again a GRINSP model showing incomplete chemistry, with $PAD=0.88$. The predicted framework is built up from corner-sharing octahedra with two quite different sizes: $[Ca_4Al_7F_{33}]$, cubic, and $a=10.876 \text{ \AA}$. The actual compound is $Na_4Ca_4Al_7F_{33}$ (Hemon and Courbion, 1990) and $a=10.781 \text{ \AA}$. By a search with chemical restraints (Ca/Al/F), the virtual model comes in fifth position (Figure 3), after 4 PDF-2 correct entries (having more or less the correct chemical formula), if the maximum angle is limited to $30^\circ 2\theta$. With a percentage of absolute difference between observed and predicted cell parameters close to 1%, large discrepancies in peak positions occur already at relatively small diffracting angles ($30^\circ 2\theta$).

D. Example 4: Zeolite–Mordenite

Most natural zeolites are not exactly SiO_2 ; even if calcined, they may contain additional cations, water molecules, etc., contrary to the hypothetical structures in the PCOD/P2D2-1. As a consequence, the discrepancies on the intensi-

ties can be large, as well as on the cell parameters. In the chosen example (Figure 4), a natural zeolite is identified as being mordenite from nine entries into the PDF-2, without any chemical restraint during the search. Next is the PCOD/P2D2-1 entry, and more PDF-2 mordenites are even listed after. Restricting EVA to the user database alone would have thus ranked the correct P2D2-1 entry first.

E. Discussion

Two main problems will obviously lead to difficulties in identification by search-match process based on a database of predicted crystal structures: inaccuracies in the predicted cell parameters, introducing discrepancies in the peak positions, and incomplete chemistry of the models, influencing the peak intensities.

Their cumulative effects will eventually be disastrous if $PAD > 2\%$, even if tuning the cell parameters is allowed in EVA in order to improve the fit (though not done for the above screen shots). However, these four cases show that identification may succeed satisfactorily if the chemistry is

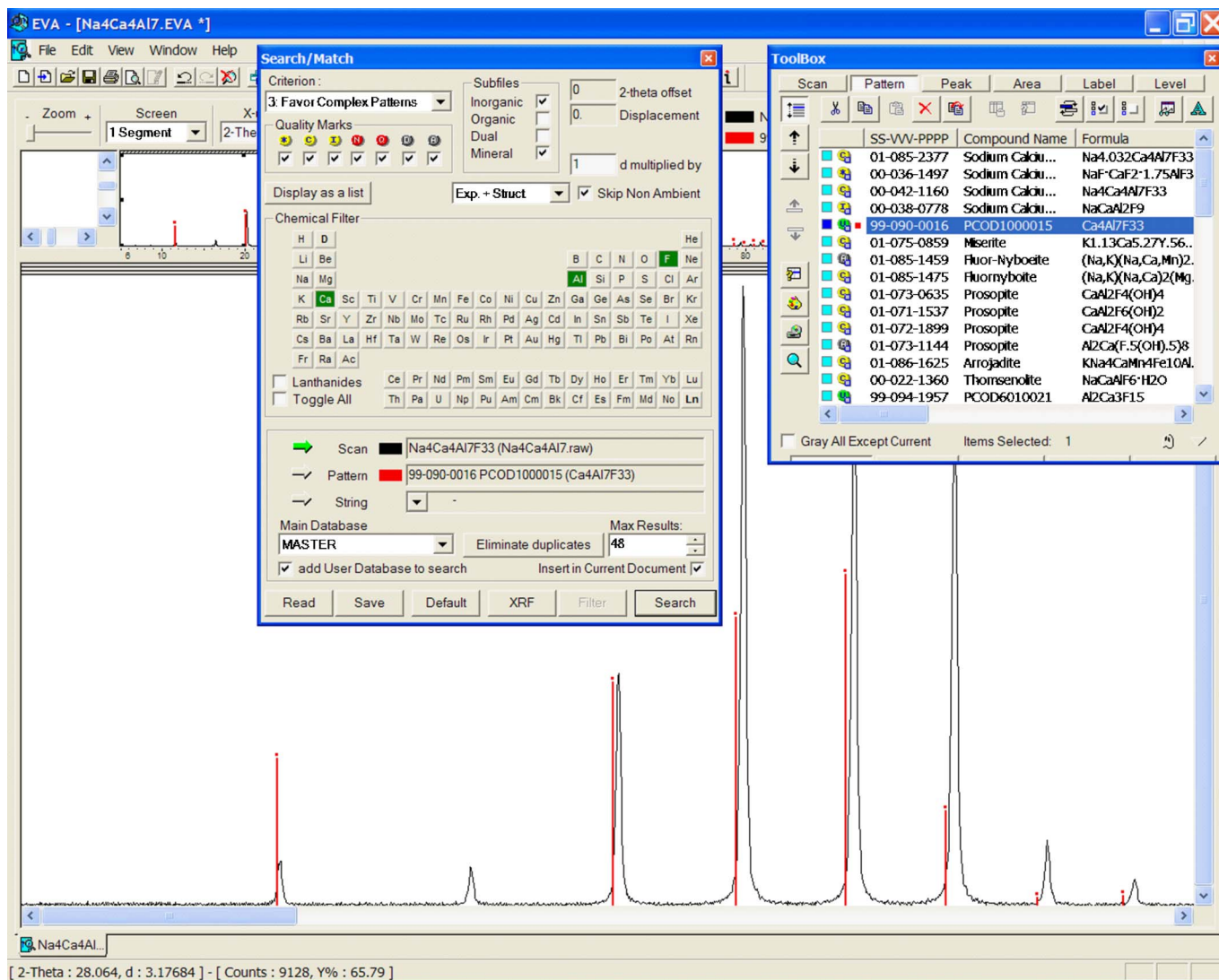


Figure 3. (Color online) Identification of $\text{Na}_4\text{Ca}_4\text{Al}_7\text{F}_{33}$ by chemically restraining the search to compounds containing Ca/Al/F in the PDF-2 and the P2D2-1.

restrained adequately during the search and if the averaged difference in cell parameters is smaller than 1%. Because cell parameter discrepancies have less sensitive effects on peak positions at low diffracting angles, for increasing the chances of success it is suggested to limit the maximum angle on the experimental diffraction pattern to 40 or even $30^\circ 2\theta$ during the search. The crystal structures of $\tau\text{-AlF}_3$ and $\text{K}_2\text{TiSi}_3\text{O}_9 \cdot \text{H}_2\text{O}$ were determined from powder data (because of the absence of suitable single crystals). Thanks to the P2D2-1 combined with search-match software, these structures would have been probably determined with much less difficulty directly at the preliminary identification stage, even before any indexing. With luck, the P2D2-1 contains a lot of to-be-discovered new crystal structures with open frameworks (titanosilicates, zeolites, etc.), which could be of interest to the microporous and metal organic framework research communities. How is it possible to attain frequently an accuracy better than 1% by using a naive cost function limited to the verification of ideal interatomic distances? The compounds predicted by GRINSP all respect most of Pauling's rules (Pauling, 1929), those stipulating vertex-connected polyhedral frameworks over edge- and face-sharing and the

parsimony of different environments of a given ion in a crystal (Meden, 2006). Attempting inorganic CSP in the larger land of edge- and face-sharing polyhedra, mixed together with corner-sharing, would require more precautions and more elaborate cost functions. Whether the PCOD contains many to-be-discovered phases and how many are useless structure candidates are hard to say. The fact is that, in spite of the difficulties to check these more than 60 000 structures against known structures in the ICSD, a large number of the hypothetical titanosilicates were found to have real counterparts (Le Bail, 2007a, 2007b) or to correspond to isostructural compounds. The aluminum fluoride series was more thoroughly investigated by DFT analysis and many virtual AlF_3 were concluded to be viable polymorphs still undiscovered, whereas the known ones were in the list of the GRINSP productions (Le Bail and Calvayrac, 2006). It is thus expected that the PCOD/P2D2-1 may help sometimes. Regular updates will be performed so that the content should increase soon in quantity and hopefully in quality with required cell parameter accuracy better than 1% as a target.

The PCOD is available and searchable at <http://>

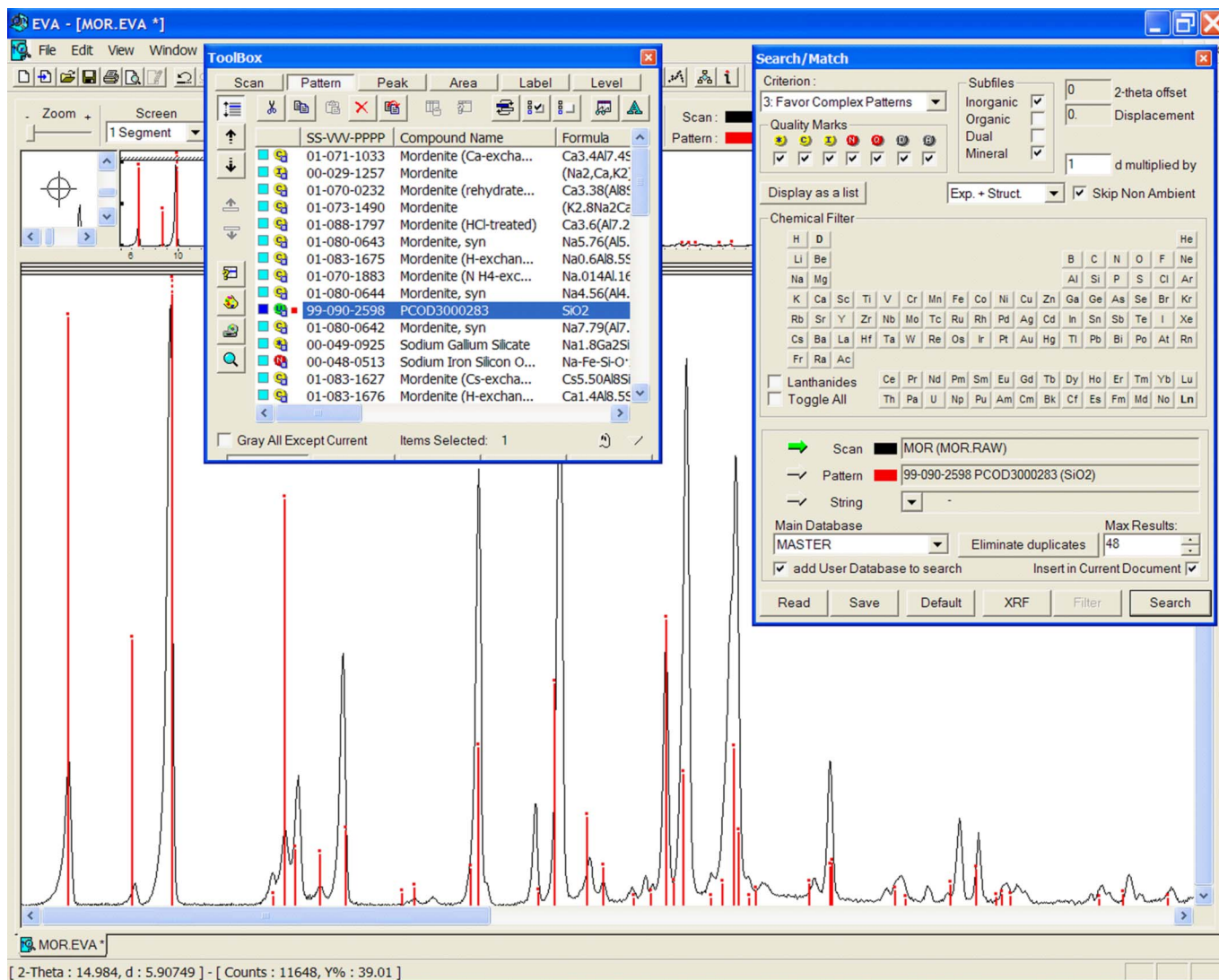


Figure 4. (Color online) Identification of a zeolite as corresponding to mordenite, one of the P2D2-1 entries matching in between all the numerous PDF-2 ones.

www.crystallography.net/pcod, where it is also possible to deposit the CIFs of new predicted crystal structures and to get the P2D2-1.

V. CONCLUSION

Prediction seems now a marginal way to solve crystal structures unsolved by the classical single crystal or powder methodologies. Nevertheless, the whole potential of the prediction approach is much broader than crystal structure solution alone. An exact theory of materials is one which would allow the full prediction of any possible crystal structure in any physical conditions in the universe and would allow for the prediction of the physical properties as well. Added to some progress in the prediction of the synthesis conditions, this would start a new era for research, overcoming the current general reliance on serendipity and providence. The PCOD and P2D2-1 databases are two very tiny steps in that challenging direction. With the search-match working effi-

ciently, the frontiers become even fuzzier between crystal structure prediction and structure determination by powder diffractometry.

- Allen, F. H. (2002). "The Cambridge Structural Database: A quarter of a million crystal structures and rising." *Acta Crystallogr., Sect. B: Struct. Sci.* **58**, 380–388.
- Bergmann, J., Le Bail, A., Shirley, R., and Zlokazov, V. (2004). "Renewed interest in powder diffraction data indexing." *Z. Kristallogr.* **219**, 783–790.
- Burla, M. C., Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G., and Siliqi, D. (2006). "Use of Patterson-based methods automatically to determine the structures of heavy-atom-containing proteins with up to 6000 non-hydrogen atoms in the asymmetric unit." *J. Appl. Crystallogr.* **39**, 728–734.
- Catlow, C. R. A. (1997). *Computer Modelling in Inorganic Crystallography* (Academic Press, London).
- Catlow, C. R. A., Gale, J. D., and Grimes, R. W. (1993). "Recent computational studies in solid state chemistry." *J. Solid State Chem.* **106**, 13–26.
- Černý, R. and Favre-Nicolin, V. (2007). "Direct space methods of structure determination from powder diffraction: principles, guidelines and perspectives." *Z. Kristallogr.* **222**, 105–113.
- Dadachov, M. S. and Le Bail, A. (1997). "Structure of zeolitic

- $K_2TiSi_3O_9 \cdot H_2O$ determined *ab initio* from powder diffraction data," Eur. J. Solid State Inorg. Chem. **34**, 381–390.
- Day, G. M., Motherwell, W. D. S., Ammon, H. L., Boerrigter, S. X. M., Della Valle, R. G., Venuti, E., Dzyabchenko, A., Dunitz, J. D., Schweizer, B., van Eijck, B. P., Erk, P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Leusen, F. J. J., Liang, C., Pantelides, C. C., Karamertzanis, P. G., Price, S. L., Lewis, T. C., Nowell, H., Torrisi, A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U., and Verwer, P. (2005). "A third blind test of crystal structure prediction," Acta Crystallogr., Sect. B: Struct. Sci. **61**, 511–527.
- Férey, G., Mellot-Draznieks, C., Serre, C., Millange, F., Dutour, J., Surblé, S., and Margiolaki, I. (2005). "A chromium terephthalate-based solid with unusually large pore volumes and surface area," Science **309**, 2040–2042.
- Férey, G., Serre, C., Mellot-Draznieks, C., Millange, F., Surblé, S., Dutour, J., and Margiolaki, I. (2004). "A hybrid solid with giant pores prepared by a combination of targeted chemistry, simulation, and powder diffraction," Angew. Chem., Int. Ed. **43**, 6296–6301.
- Fischer, C. C., Tibbetts, K. J., Morgan, D., and Ceder, G. (2006). "Predicting crystal structure by merging data mining with quantum mechanics," Nat. Mater. **5**, 641–646.
- Foster, M. D., Friedrichs, O. D., Bell, R. G., Paz, F. A. A., and Klinowski, J. (2003). "Structural evaluation of systematically enumerated hypothetical uninodal zeolites," Angew. Chem., Int. Ed. **42**, 3896–3899.
- Foster, M. D. and Treacy, M. M. J. (2003). Hypothetical Zeolites Database (<http://www.hypotheticalzeolites.net>).
- Gale, J. D. (1997). "GULP: A computer program for the symmetry-adapted simulation of solids," J. Chem. Soc., Faraday Trans. **93**, 629–637.
- Gavezzotti, A. (1994). "Are crystal structures predictable?," Acc. Chem. Res. **27**, 309–314.
- Hemon, A. and Courbion, G. (1990). "The NaF-CaF₂-AlF₃ system: Structures of β -NaCaAlF₆ and Na₄Ca₄Al₇F₃₃," J. Solid State Chem. **84**, 153–164.
- Hofmann, D. W. M. and Kuleshova, L. (2005). "New similarity index for crystal structure determination from X-ray powder diagrams," J. Appl. Crystallogr. **38**, 861–866.
- Le Bail, A. (2003). PCOD: Predicted Crystallography Open Database (<http://www.crystallography.net/pcod>).
- Le Bail, A. (2005). "Inorganic structure prediction with GRINSP," J. Appl. Crystallogr. **38**, 389–395.
- Le Bail, A. (2007a). "Predicted corner-sharing titanium silicates," Z. Kristallogr. Suppl. **26**, 202–208.
- Le Bail, A. (2007b). "Inorganic structure prediction: Too much and not enough," Solid State Phenom. **130**, 1–6.
- Le Bail, A. and Calvayrac, F. (2006). "Hypothetical AlF₃ crystal structures," J. Solid State Chem. **179**, 3159–3166.
- Le Bail, A., Fourquet, J. L., and Bentrup, U. (1992). "7-AlF₃: Crystal structure determination from X-ray powder diffraction data. A new MX₃ corner-sharing octahedra 3D network," J. Solid State Chem. **100**, 151–159.
- Le Meins, J.-M., Cranswick, L. M. D., and Le Bail, A. (2003). "Results and conclusions of the internet based search/match round robin 2002," Powder Diffr. **18**, 106–113.
- Lufaso, M. W. and Woodward, P. M. (2001). "Prediction of the crystal structures of perovskites using the software SPuDS," Acta Crystallogr., Sect. B: Struct. Sci. **57**, 725–738.
- Meden, A. (2006). "Inorganic crystal structure prediction – a dream coming true?" Acta Chim. Slov. **53**, 148–152.
- Mellot-Draznieks, C. and Férey, G. (2005). "Assembling molecular species into 3D frameworks: Computational design and structure solution of hybrid materials," Prog. Solid State Chem. **33**, 187–197.
- Mellot-Draznieks, C., Girard, S., Férey, G., Schön, J. C., Cancarevic, Z., and Jansen, M. (2002). "Computational design and prediction of interesting not-yet-synthesized structures of inorganic materials by using building unit concepts," Chem.-Eur. J. **8**, 4102–4113.
- Mellot-Draznieks, C., Newsam, J. M., Gorman, A. M., Freeman, C. M., and Férey, G. (2000). "De novo prediction of inorganic structures developed through automated assembly of secondary building units (AASBU method)," Angew. Chem., Int. Ed. **39**, 2270–2275.
- Milman, V. and Winkler, B. (1999). "Ab initio modeling in crystallography," Int. J. Inorg. Mater. **1**, 273–279.
- Molecular Simulations (2000). Cerius², Version 4.2 (Computer Software), Accelrys Software Inc., Cambridge, United Kingdom.
- Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Dzyabchenko, A., Erk, P., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Lommerse, J. P. M., Mooij, W. T. M., Price, S. L., Scheraga, H., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P., and Williams, D. E. (2002). "Crystal structure prediction of small organic molecules: A second blind test," Acta Crystallogr., Sect. B: Struct. Sci. **58**, 647–661.
- Panina, N., Leusen, F. J. J., Janssen, F. F. B. J., Verwer, P., Meekes, H., Vlieg, E., and Deroover, G. (2007). "Crystal structure prediction of organic pigments: Quinacridone as an example," J. Appl. Crystallogr. **40**, 105–114.
- Pauling, L. (1929). "The principles determining the structure of complex ionic crystals," J. Am. Chem. Soc. **51**, 1010–1026.
- Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. A., and Joannopoulos, J. D. (1992). "Iterative minimization techniques for *ab initio* total-energy calculations: Molecular dynamics and conjugate gradients," Rev. Mod. Phys. **64**, 1045–1097.
- Rietveld, H. M. (1969). "A profile refinement method for nuclear and magnetic structures," J. Appl. Crystallogr. **2**, 65–71.
- Schmidt, M. U., Ermrich, M., and Dinnebier, R. E. (2005). "Determination of the structure of the violet pigment C₂₂H₁₂Cl₂N₆O₄ from a non-indexed X-ray powder diagram," Acta Crystallogr., Sect. B: Struct. Sci. **61**, 37–45.
- Schmidt, M. U., Hofmann, D. W. M., Buchsbaum, C., and Metz, H. J. (2006). "Crystal structure of pigment Red 170 and derivatives, as determined by X-ray powder diffraction," Angew. Chem., Int. Ed. **45**, 1313–1317.
- Schön, J. C. and Jansen, M. (2001a). "Determination, prediction, and understanding of structures, using the energy landscapes of chemical systems – Part I," Z. Kristallogr. **216**, 307–325.
- Schön, J. C. and Jansen, M. (2001b). "Determination, prediction, and understanding of structures, using the energy landscapes of chemical systems – Part II," Z. Kristallogr. **216**, 361–383.
- Treacy, M. M. J., Rivin, I., Balkovsky, E., Randall, K. H., and Foster, M. D. (2004). "Enumeration of periodic tetrahedral frameworks. II. Polynodal graphs," Microporous Mesoporous Mater. **74**, 121–132.
- Wevers, M. A. C., Schön, J. C., and Jansen, M. (1998). "Determination of structure candidates of simple crystalline AB₂ systems," J. Solid State Chem. **136**, 233–246.
- Winkler, B., Knorr, K., and Milman, V. (2003). "Prediction of the structure of LaF₃ at high pressures," J. Alloys Compd. **349**, 111–113.
- Woodley, S. M. (2004). "Prediction of crystal structures using evolutionary algorithms and related techniques," in *Application of Evolutionary Computation in Chemistry*, edited by D. M. P. Mingos and R. L. Johnston (Springer-Verlag, Berlin), Vol. 110, pp. 95–132.
- Yaghi, O. M., O’Keeffe, M., Ockwig, N. W., Chae, H. K., Eddaoudi, M., and Kim, J. (2003). "Reticular synthesis and the design of new materials," Nature (London) **423**, 705–714.