Crystal Structure Prediction

DOI: 10.1002/anie.201101530

Crystal Structure Prediction—Dawn of a New Era

*Christian W. Lehmann**

crystal structure prediction · density functional calculations · polymorphism · structure elucidation

Every year thousands of new crystal structures are determined and many of these are molecular crystal structures. However, while an enormous amount of structural data is available, it has proven rather difficult to predict molecular crystal structures given only the structural formula. As with all chemical transformations, a change in energy is the driving force underlying the formation of any crystal, and both kinetic as well as thermodynamic contributions are important. Molecular crystals are held together primarily by dispersion forces which are comparatively weak and notoriously difficult to calculate accurately. In 1994 in a seminal paper entitled *"Are Crystal Structures Predictable?"* Gavezotti stated a clear "No" and pointed out that the small energy differences between hypothetical crystal structures are the limiting factor for successful crystal structure predictions.[1]

Since then, this dogma has been challenged numerous times, most prominently by a series of blind tests organized by the Cambridge Crystallographic Data Centre (CCDC). So far a total of five such blind tests have been organized and the crystal structure predictions have become more accurate, despite increased challenges posed by the organizers.[2] For these tests a small number of organic compounds are selected with hitherto unpublished crystal structures. The participants are asked to propose three candidate crystal structures for each compound which are matched against the secret experimental crystal structures. For the latest series, which was concluded in 2010, the results have been submitted and analyzed but not yet published. This study included six categories of increasing complexity and, for the first time, it also included systems with known polymorphism and more flexible molecules.[3]

Various strategies are pursued in the field of crystal structure prediction in order to generate hypothetical crystal structures and most importantly in order to identify and rank likely structures.[4] While one set of methods relies on calculating the overall energy at different approximations, another, more empirical strategy employed by Desiraju and others utilizes the concept of supramolecular synthons. In this approach the intermolecular interaction energies are considered implicitly based on frequently occurring motifs.[5] While chemical intuition and experience are essential ingredients for crystal engineering, the application of the Cambridge Structural Database relies more rigorously on historical facts (i.e. an archive of experimentally determined crystal structure) and statistical tools. This approach, in combination with a distributed multipole model for describing atomic charges, has been shown to work for rigid aromatic molecules.[6] The majority of strategies, however, are based on intermolecular energies alone. The results of the CCDC blind test of 2007 illustrate aptly the continued difficulties. Only three of the participating research groups predicted two of a possible four crystal structures correctly, and only one group correctly predicted all four structures.[2d]

One aspect that demonstrates the subtleties of crystal structure prediction is polymorphism.[7] The observation that the same chemical entity (molecule) is able to crystallize with more than one spatial arrangement of the atoms as a result of a change, sometimes only minimal, in the crystallization conditions, shows that several local minima can exist on the energy landscape. Leusen et al. recently reported the successful calculation of all three known crystal forms of molecule VI (6-amino-2-phenylsulfonylimino-1,2-dihydropyridine).[8] The significant advance is the successful ranking of these crystal structures on an energy scale and the clear separation in energy of these known crystal forms (polymorphs) from other hypothetical crystal structures. The computational approach followed by Leusen et al. is based on a combination of dispersion-corrected density functional theory (DFT(d)) and molecular mechanics. In order to sample the parameter space efficiently and to produce a subset of candidate structures a tailor-made force field is employed. Resulting candidate structures are minimized subsequently at the DFT(d) level. The power of the method has been demonstrated impressively;[9] however, the method is computationally rather expensive. The total CPU time used for DFT(d) method in the 2007 blind test was about 280 000 h, while other methods required only up to 5000 h. These computational costs might seem expensive, maybe even prohibitively expensive for an academic environment, but compared to the costs of developing a pharmaceutical ingredient into a marketable solid form the computational cost is hardly a limiting factor.

Polymorphism not only exemplifies the finely tuned energetics of crystal structures, it is also of major importance for the actual manufacture of molecular solids, for example organic pigments and active pharmaceutical ingredients. The

[*] Prof. C. W. Lehmann
Max-Planck-Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr (Germany)
Fax: (+49) 208-306-2989
E-mail: lehmann@kofo.mpg.de

latter compounds in particular are rich in unexpected effects that control the bioavailability of drugs (for example Ritonavir) and in legal battles over intellectual property rights (for example Ranitidine hydrochloride).[10] Being able to predict conclusively all possible polymorphs of a given molecule, irrespective of how any or all of these crystal forms are made, would support pharmaceutical companies in their development of reliable and safe products.

As Leusen et al. point out, the successful prediction of all likely polymorphs of a given organic molecule is no guarantee that these polymorphs can actually be crystallized, nor does it provide a recipe for carrying out the crystallization. Presently it might not be feasible to perform crystallization experiments in silico; however, not too long ago the same was said about calculating reaction pathways.

Received: March 2, 2011
Published online: May 5, 2011

[1] A. Gavezzotti, *Acc. Chem. Res.* **1994**, *27*, 309 – 314.
[2] a) J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer, D. E. Williams, *Acta Crystallogr. Sect. B* **2000**, *56*, 697 – 714; b) W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer, D. E. Williams, *Acta Crystallogr. Sect. B* **2002**, *58*, 647 – 661; c) G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, P. Verwer, *Acta Crystallogr. Sect. B* **2005**, *61*, 511 – 527; d) G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf, B. Schweizer, *Acta Crystallogr. Sect. B* **2009**, *65*, 107 – 125.
[3] D. Bardwell, *Crystalline—CCDC Newsletter*, November **2010** http://www.ccdc.cam.ac.uk/news/newsletter/nov_10.pdf.
[4] G. M. Day, *Crystallogr. Rev.* **2011**, *17*, 3 – 52.
[5] J. A. R. P. Sarma, G. R. Desiraju, *Cryst. Growth Des.* **2002**, *2*, 93 – 100.
[6] S. L. Price, *Acc. Chem. Res.* **2009**, *42*, 117 – 126.
[7] J. Bernstein, *Polymorphism in Molecular Crystals*, Clarendon, Oxford, **2002**.
[8] H. C. S. Chan, J. Kendrick, F. J. J. Leusen, *Angew. Chem. Int. Ed.* **2011**, *50*, 2979 – 2981.
[9] M. A. Neumann, F. J. J. Leusen, J. Kendrick, *Angew. Chem.* **2008**, *120*, 2461 – 2464; *Angew. Chem. Int. Ed.* **2008**, *47*, 2427 – 2430.
[10] *Polymorphism in the Pharmaceutical Industry* (Ed.: R. Hilfiker), Wiley-VCH, Weinheim, **2006**.

