Torsten Stelzer (Autor)
Joachim Ulrich (Autor)

# BIWIC 2010

*17th International Workshop on Industrial Crystallization*


https://cuvillier.de/de/shop/publications/673

# A decade of blind tests in crystal structure prediction

F. J. J. Leusen

School of Life Sciences, University of Bradford, Bradford, BD7 1DP, United Kingdom
f.j.j.leusen@bradford.ac.uk

*The goal of predicting the crystal structures of an organic molecule from its molecular structure alone has attracted considerable interest from industry and academia. The difficulty of the task is demonstrated by the regular 'Crystal Structure Prediction Blind Test' organised by the Cambridge Crystallographic Data Centre in 1999, 2001, 2004, 2007 and 2010. In this contribution, the results of the first four Blind Tests will be reviewed, and some initial results of the most recent event will be presented.*

## 1. Introduction

The solid state properties of organic materials are of paramount importance for a wide range of industrial applications, including pharmaceuticals, agrochemicals, pigments, explosives and food stuff. These properties depend not only on molecular structure, but also on how the molecules pack together in the crystalline environment. The phenomenon of polymorphism, which is the ability of a molecule to crystallise in more than one distinct crystal structure, complicates matters further because the physico-chemical properties of different polymorphs can vary substantially.

The ability to predict the crystal structures of an organic compound has been an active research topic for several decades. Crystal structure prediction can be applied to deduce the crystal structure of a compound for which only low quality experimental X-ray powder diffraction data is available, to predict the crystal structure of a compound prior to experimentation (e.g., in pigment design), or to investigate polymorphism. The latter application is of particular interest in the pharmaceutical industry, where the appearance of a new polymorphic form after the drug delivery system for a new pharmaceutical has been decided can cause major problems. Accurate and reliable crystal structure prediction is considered by some as the holy grail of crystal engineering. This ultimate aim has still not been achieved, although significant advances have been made in recent years.

To test the state-of-the-art in crystal structure prediction methodologies, a series of Blind Tests have been organised by the Cambridge Crystallographic Data Centre. In each Blind Test, a number of organic compounds for which the experimental crystal structures have been determined accurately are selected by an independent referee. Blind Test participants are provided with the molecular structures of the test compounds whilst the experimental crystal structures are kept hidden. Participants have a limited period of time to submit up to three predictions for each of the compounds before the experimental data is disclosed and the predictions are evaluated.

This contribution will illustrate the progress in crystal structure prediction by reviewing the original results of the first four Blind Tests, organised in 1999 [1], 2001 [2], 2004 [3] and 2007 [4]. The results of a recent study [5], in which the predictions made in the first three Blind Tests were re-evaluated using the method that predicted all four structures correctly in the 2007 Blind Test [6], will also be discussed. Finally, some initial results of the latest Blind Test, conducted in 2010, will be presented.

## 2. Overview of Methods

Crystal structure prediction is dominated by two main problems. The first, mathematical, problem involves searching for all possible crystal packing alternatives for a given molecule. This search needs to be reliable in order to not miss any important structures. The second, physical, problem is posed by the necessity to rank all potential polymorphs according to their stability. It is assumed that only the most stable crystal structures will be observed experimentally. Here, accuracy is of key importance because the stability differences between polymorphs are small. Both problems can be tackled in a number of ways, and the Blind Test participants have used a wide variety of methods to make their predictions [1-4]. However, all successful predictions were made using methods which approximated polymorphic stability by the calculation of lattice energies, the majority by means of molecular mechanics methods.

In the 2007 Blind Test, a new method emerged that uses solid state quantum mechanics to calculate the relative stabilities of potential polymorphs [6, 7]. The essential, new feature of the approach is a hybrid method for the calculation of lattice energies that combines quantum mechanical, density functional theory simulations with an empirical van der Waals correction [7]. Since the hybrid method is too demanding computationally to be used during the search for crystal packing alternatives, it is first used to parameterise a tailor made, non-transferable, force field for each compound [8]. Potential crystal structures are then generated using this force field, and finally the most promising candidates are fully optimised using the hybrid method.

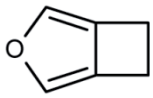
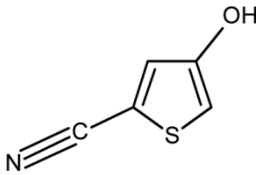
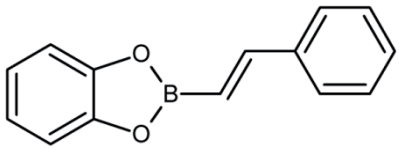
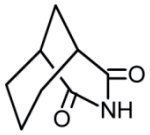
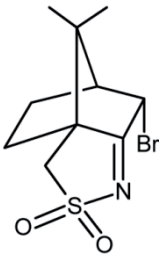
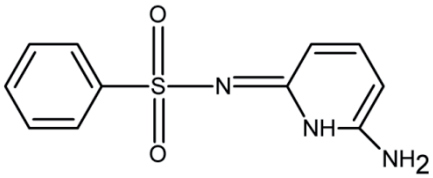
## 3. Results and Discussion

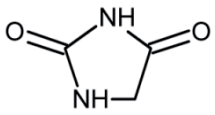
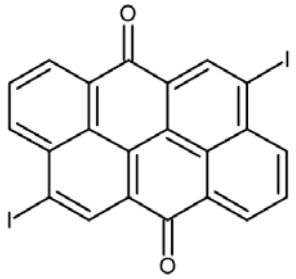
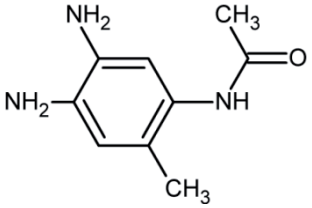
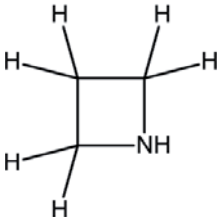
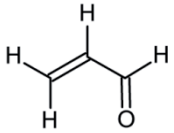
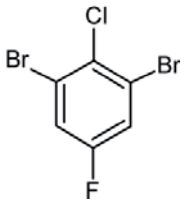
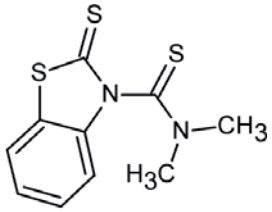
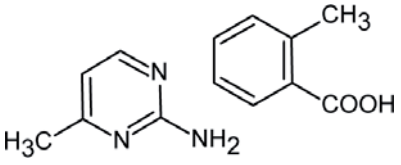
In the first three Blind Tests (1999 [1], 2001 [2] and 2004 [3]), a limited number of successful predictions were reported, all of which were achieved using force field methods, where the intermolecular and intramolecular forces are represented by analytical functions. In the 2007 Blind Test, the new hybrid method was successfully applied to all four test compounds, including a co-crystal [4, 6]. The molecular structures and prediction results are reported in Tab. 1.

To further validate the hybrid method, it was applied retrospectively to the first three Blind Tests to recalculate the relative stability of all submitted and experimental crystal structures. Considering the first four Blind Tests, the hybrid method predicted the experimentally observed crystal structure as the most stable structure for 13 out of 15 compounds [5]. Additional polymorphs of molecules IV and VI, which were discovered after the 2001 Blind Test took place, were ranked as the second most stable structure for both molecules. Furthermore, it is predicted that a new polymorph of one of the test compounds (molecule V) exists under pressure [5]. These results are also presented in Tab. 1.

The fifth Blind Test, consisting of six test compounds including a salt, a hydrate and a large flexible molecule, is currently underway (see Fig. 1 for the molecular structures). Initial results will be known in September 2010 and will be previewed at this conference.

**Tab. 1:** Molecular structures and crystal structure prediction results for the first fou- Blind tests.

| Molecule | Molecular structure | # correct predictions / # participants | Rank with hybrid method | Hybrid method ΔE (kcal/mol) |
|---|---|---|---|---|
| I |  Two polymorphs | A: 4 / 11<br>B: 0 / 11 | 1<br>2 | 0.00<br>0.12 |
| II |  | 1 / 8 | 2 | 0.01 |
| III |  | 1 / 11 | 1 | 0.00 |
| IV |  Second polymorph discovered after Blind Test | A: 2 / 15<br>B: Not part of Blind Test | 1<br>2 | 0.00<br>0.13 |
| V |  | 4 / 15 | 4 | 0.36<br>new form predicted |
| VI |  Second polymorph discovered after Blind Test | A: 0 / 11<br>B: Not part of Blind Test | 1<br>2 | 0.00<br>0.22 |
| VII |  | 1 / 4 | 1 | 0.00 |

| | | | | |
|---|---|---|---|---|
| VIII |  <br>Not blind<br>(partial structure known) | 4 / 15 | 1 | 0.00 |
| IX |  | 1 / 15 | 1 | 0.00 |
| X |  | 0 / 15 | 1 | 0.00 |
| XI |  <br>Z' = 2 | 0 / 18 | 1 | 0.00 |
| XII |  | 4 / 15 | 1 | 0.00 |
| XIII |  | 4 / 14 | 1 | 0.00 |
| XIV |  | 3 / 14 | 1 | 0.00 |
| XV |  <br>Co-crystal | 2 / 12 | 1 | 0.00 |

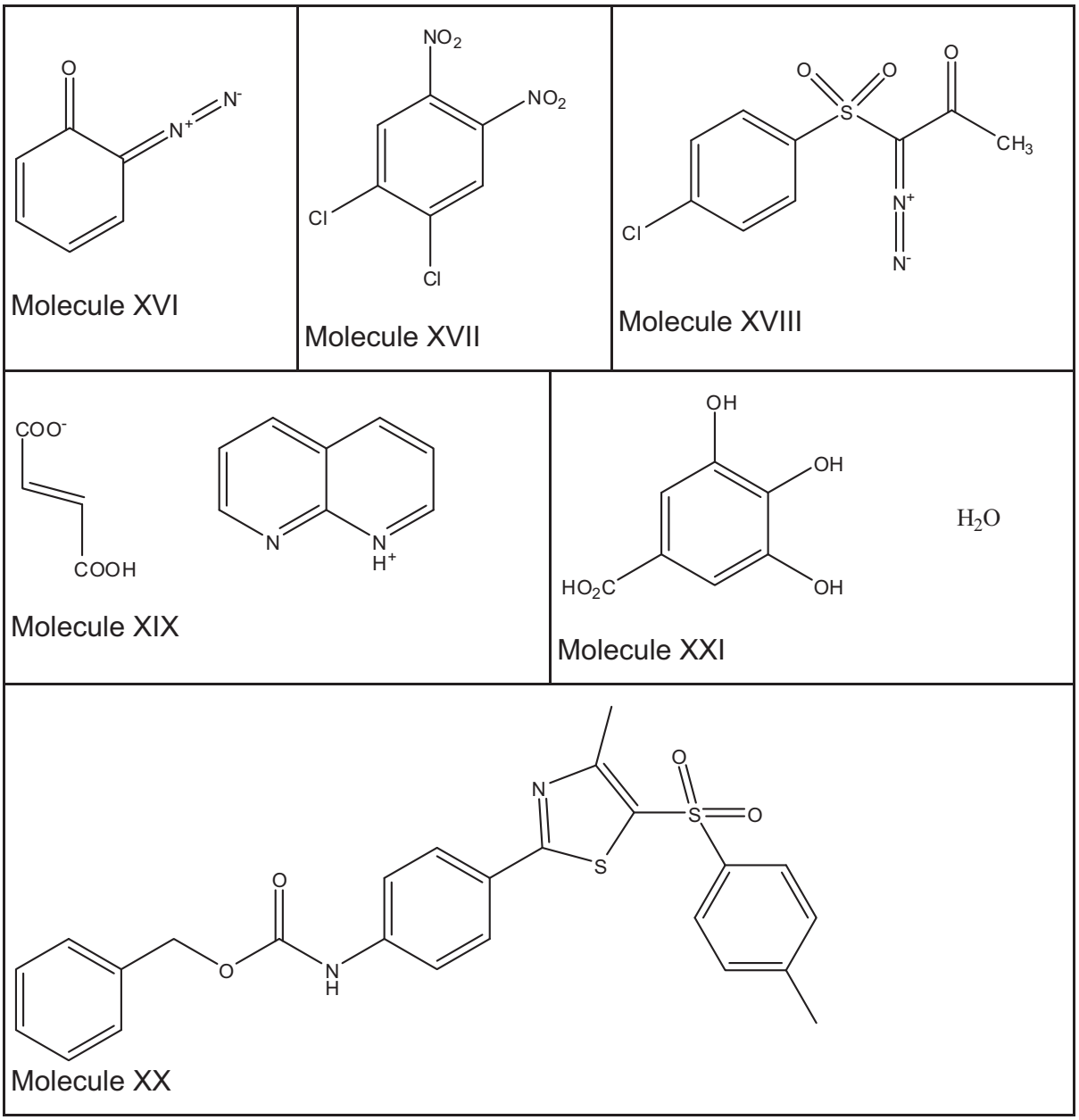



**Fig. 1:** Molecular structures of compounds selected for the 2010 Blind Test.

## 4. Conclusions

Whilst crystal structure prediction can still not be considered as a general polymorph screening tool, major advances have been made in recent years. The key to this progress has been the accurate calculation of lattice energies by means of solid state quantum mechanical methods. The solid state structures of small, semi-rigid molecules can now be routinely predicted. Larger, more flexible molecules are also becoming feasible, albeit at the expense of significant computational effort. The same applies to salts, hydrates and other systems with more than one independent molecule in the asymmetric unit, although ionic interactions may pose problems.

Crystallisation is a kinetic process, and true prediction of polymorphic behaviour can only be achieved by considering entropic and zero-point energy contributions. Such calculations are feasible but extremely demanding and are therefore not routinely

applied. Future work will also need to consider the effect of solvent in polymorph selection, however, the accurate simulation of solvent interactions and their impact on the relative stability of polymorphs is currently not yet feasible.

References

[1] J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, 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: *A test of crystal structure prediction of small organic molecules*, Acta Cryst. B, 56 (2000) 697-714.

[2] 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: *Crystal structure prediction of small organic molecules: a second blind test*, Acta Cryst. B, 58 (2002) 647-661.

[3] 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: *A third blind test of crystal structure prediction*, Acta Cryst. B, 61 (2005) 511-527.

[4] 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: *Significant progress in predicting the crystal structures of small organic molecules – a report on the fourth blind test*, Acta Cryst. B, 65 (2009) 107-125.

[5] A. Asmadi, M. A. Neumann, J. Kendrick, P. Girard, M.-A. Perrin, F. J. J. Leusen: *Revisiting the blind tests in crystal structure prediction: accurate energy ranking of molecular crystals*, J. Phys. Chem. B, 113 (2009) 16303-16313.

[6] M. A. Neumann, F. J. J. Leusen, J. Kendrick: *A major advance in crystal structure prediction*, Angew. Chem. Int. Ed., 47 (2008) 2427-2430.

[7] M. A. Neumann, M.-A. Perrin: Energy ranking of molecular crystals using density functional theory calculations and an empirical van der Waals correction, J. Phys. Chem. B, 109 (2005) 15531-15541.

[8] M. A. Neumann: *Tailor-made force fields for crystal-structure prediction*, J. Phys. Chem. B, 112 (2008) 9810–9829.