Chemical Engineering Science 121 (2015) 60–76

Contents lists available at ScienceDirect

Chemical Engineering Science

journal homepage: www.elsevier.com/locate/ces

Prediction of the crystal structures of axitinib, a polymorphic pharmaceutical molecule

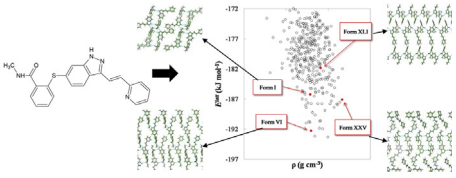
Manolis Vasileiadis, Constantinos C. Pantelides, Claire S. Adjiman *

Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, United Kingdom

HIGHLIGHTS

- First ab initio prediction of the polymorphs of axitinib, a challenging pharmaceutical molecule.
- Successful application of a systematic procedure to identify polymorphs with one molecule in asymmetric unit.
- Analysis of strengths and weaknesses of state of the art for crystal structure prediction.
- Identification of future research needs and priorities in this area.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 26 May 2014
Received in revised form
11 August 2014
Accepted 30 August 2014
Available online 10 September 2014

Keywords:
Polymorph
Lattice energy
CrystalPredictor
CrystalOptimizer
Local approximate models

ABSTRACT

Organic molecules can crystallize in multiple structures or polymorphs, yielding crystals with very different physical and mechanical properties. The prediction of the polymorphs that may appear in nature is a challenge with great potential benefits for the development of new products and processes. A multistage crystal structure prediction (CSP) methodology is applied to axitinib, a pharmaceutical molecule with significant polymorphism arising from molecular flexibility. The CSP study is focused on those polymorphs with one molecule in the asymmetric unit. The approach successfully identifies all four known polymorphs within this class, as well as a large number of other low-energy structures. The important role of conformational flexibility is highlighted. The performance of the approach is discussed in terms of both the quality of the results and various algorithmic and computational aspects, and some key priorities for further work in this area are identified.

1. Introduction

Organic crystals are a key component of the formulated products that are manufactured in many industrial sectors including pharmaceuticals (Storey and Ymén, 2011), agrochemicals, foods, paints, and explosives. The efficacy, stability and other end-use properties of such products are largely influenced by the precise structure of the organic crystals because the molecular packing arrangement affects numerous physical properties such as color, mechanical strength, flowability and solubility, to name but a few (Hilfiker, 2006).

In view of the importance of crystal structure, the propensity of many organic molecules to crystallize readily in multiple metastable structures ("polymorphs") (Brog et al., 2013; Cruz-Cabeza and Bernstein, 2014) creates significant challenges in many aspects of product development and manufacturing. A well-known example of the problems that can arise as a result of polymorphism is that of ritonavir (Bauer et al., 2001), an active pharmaceutical ingredient (API) marketed as a HIV drug by Abbott Laboratories from 1996. In 1998, a previously unknown form, Form II, appeared, and it became impossible to revert to the production of Form I. Form II was found to be more stable than Form I, with a

* Corresponding author. Tel.: +44 20 7594 6638.
E-mail address: c.adjiman@imperial.ac.uk (C.S. Adjiman).

http://dx.doi.org/10.1016/j.ces.2014.08.058

significantly lower solubility. The product had to be recalled from the market, leading to an interruption of supply, and was eventually re-developed as a liquid formulation. Further investigation revealed three further forms of ritonavir (Morissette et al., 2003). The existence of polymorphs also poses intellectual property challenges, as patent protection relates to the form of the product. For example, the crystal structures of cefdinir, a drug molecule with at least five polymorphs, have been the subject of multiple patents and of prolonged legal battles (Cabri et al., 2007).

The magnitude of the risks arising from insufficient knowledge of polymorphism has motivated increasing investment in polymorph screening, the experimental investigation of the so-called polymorphic landscape of organic molecules (Aaltonen et al., 2009; Newman, 2013). This has been complemented by computational crystal structure prediction (CSP) methodologies aiming to identify possible crystal structures with little or no experimental input. While it has long been clear that achieving this goal would require a very significant research effort (Gavezzotti, 1994), CSP is increasingly used in combination with experimental screening (Price, in press). The blind tests organized by the Cambridge Crystallographic Data Centre since 1999 (Lommerse et al., 2000) provide a useful series of snapshots of the state-of-the-art and of the progress made in the field. In each blind test, participants are asked to predict the most stable crystal structure for a handful of molecules, salts or co-crystals of varying complexity. The degree of difficulty of each system depends on the number of molecules it contains, the types and number of atoms, the presence of charged species and the flexibility of the molecules. Of particular note in previous blind tests are two milestones: the consistent success achieved with the GRACE approach by Neumann, Leusen, Kendrick, in the fourth blind test (Day et al., 2009) and by Neumann, Leusen, Kendrick, and van de Streek in the fifth blind test (Bardwell et al., 2011), in predicting the polymorphs of small molecules (Kendrick et al., 2011); and the successful prediction in the fifth blind test, by two groups (Kazantsev et al., 2011b), of the most stable structure of "Molecule XX" (benzyl-(4-(4-methyl-5-(p-tolylsulfonyl)-1,3-thiazol-2-yl)phenyl)carbamate), a molecule whose structure, size and flexibility (see Fig. 1g) are representative of those of pharmaceutical compounds.

As a result of the increasing reliability of CSP (Day, 2011; Price, in press), several promising applications to industrially-relevant compounds have been reported in the literature, focusing on the identification of known and potential polymorphs. In the area of pharmaceuticals, these have included studies of some of the compounds shown in Fig. 1, namely (a) naproxen (Braun et al., 2011), (b) GlaxoSmithKline's molecule GSK269984B (Ismail et al., 2013), (c) Pfizer's crizotinib (Abramov, 2013), (d) a melatonin agonist (Kendrick et al., 2013), (e) Eli Lilly's olanzapine (Bhardwaj et al., 2013) and (f) Eli Lilly's tazofelone (Price et al., in press). In these different cases, crystal energy landscapes, in which every putative crystal structure is characterized in terms of its energy and density, were generated. The computed crystal structures were ranked in terms of their thermodynamic stability, usually based on the predicted lattice energy, rather than the more difficult to compute Gibbs free energy.

The studies of pharmaceutical compounds reported in the literature to date have focused on "small molecule pharmaceuticals", typically with up to 10 rotatable bonds. They have generally resulted in the correct identification of all known polymorphic structures as low-energy minima on the energy landscape. The relative stability of the computed polymorphs, however, often differs from the experimental relative stability, as extrapolated to 0 K. Furthermore, many structures that have not been identified experimentally are often found as low-energy minima. This can arise for a number of reasons, including the fact that some computed structures may be found to be unstable when entropic effects are taken into account (Mooij et al., 1998; Zykova-Timan et al., 2008) and the fact that some structures may be difficult to crystallize experimentally (Price, 2013).

Despite these limitations, CSP has found several applications of practical relevance beyond the scientific goal of achieving the blind prediction of all likely polymorphs. Thus, it can be used (i) to provide reassurance that all likely polymorphs have been identified; (ii) to guide the search for further polymorphs by suggesting the specific crystal structures that might be observed, thereby helping to identify appropriate crystallization conditions (Arlin et al., 2011); (iii) to support crystal engineering by providing an understanding of the link between the motifs observed and molecular structure (Uzoh et al., 2012) or crystal composition (Habgood, 2013; Karamertzanis et al., 2009); (iv) to help crystallographers interpret data gathered on specific compounds (Baias et al., 2013; Friščić et al., 2010; Wu et al., 2013). This broad array of uses provides impetus for methodological improvements aimed at increasing the accuracy of the predictions and at broadening the range of molecules, co-crystals, salts and solvates that can be tackled in terms of size and complexity.

Several reviews have recently been published on the current state-of-the-art in CSP, covering one or more methodologies (Abramov, 2013; Atahan-Evrenk and Aspuru-Guzik, 2014; Day, 2010, 2011, 2012; Kendrick et al., 2011; Pantelides et al., 2014; Price, 2013, 2008a, 2008b, in press). Together with the papers summarizing the results of the five blind tests to date (Bardwell et al., 2011; Day et al., 2009, 2005; Lommerse et al., 2000; Motherwell et al., 2002), these provide an excellent survey of the field. In the present paper, we focus on a specific systematic approach that has been developed in our group (Pantelides et al., 2014). The algorithms on which this approach is based have been successfully used in several of the examples discussed so far (Bardwell et al., 2011; Bhardwaj et al., 2013; Ismail et al., 2013; Kazantsev et al., 2011b; Price et al., in press; Vasileiadis et al., 2012). We aim to provide an introduction to the approach and a perspective on future developments via its application to axitinib (Fig. 2), a Pfizer anti-cancer API that has been noted for its numerous crystal forms, including 5 neat ones and 66 solvates. This provides a great challenge for CSP and a fertile learning ground allowing us to assess the current status of the methodology and to identify directions for further research. In Section 2, we review previous work on the crystal structures of axitinib. In Section 3, we provide an overview of our CSP methodology, and in Section 4 we discuss its application to the prediction of polymorphism in axitinib. Section 5 discusses various key aspects of the performance of the CSP approach when applied to axitinib, aiming to draw some lessons from the results obtained. Section 6 concludes with some general remarks on the current status of CSP methodologies and their limitations, and identifies some relevant research priorities in this area.

2. Earlier work on the crystal structures of axitinib

This section provides a review of the available information on the polymorphism of axitinib. Published experimental data on the five neat polymorphs are summarized, and previous computational work is discussed.

2.1. Experimental investigations

From the point of view of crystallography, axitinib is notable because of its large number of neat polymorphs, solvates and hydrates. The 71 forms that have been reported in the literature to