An appreciation of organic solid state chemistry and challenges in the field of “molecules, materials, medicines”.

William Jones, Department of Chemistry, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

Abstract

Over many years much of what has motivated a considerable amount of research in the general area of organic solid state chemistry concerns an understanding of how the molecular packing within an organic solid affects its physical and chemical properties. Whilst the motivation for this understanding may have changed over time the fundamental issues remain the same. An understanding of the effect of crystal attributes (e.g. polymorph, habit and particle size) on chemical, physical, photophysical and electronic properties is still vital to the development of organic solids with optimised properties. While progress has been made in analytical skills (including sophisticated developments in instrumentation) along with major developments in computational techniques, there are still many challenges. This paper, with a primary focus on pharmaceuticals, provides a brief, and not comprehensive, personal overview of the progress which has been made and has yet to be made. Amongst the chemists and crystallographers who have contributed significantly to the subject is Jack Dunitz and some of his seminal papers appear in several of the issues discussed.

1. Introduction

In an article published in 2003 Jack Dunitz addressed the issue of “Are crystal structures predictable”?\(^1\) This article, and one by Gavezzotti,\(^2\) concerned the issue of developments in the field of predicting crystal structures from a knowledge of molecular structure alone. Significant strides had been made following the challenge laid down by the then Editor of Nature concerning the “continuing scandal in the physical sciences that it is still impossible to predict the crystal structure of even a simple crystal”.\(^3\) Despite the pioneering work of Kitaigorodskii and others\(^4\) and subsequent developments in our understanding of molecular recognition with regard to the process of crystallisation\(^5\) there was validity to this statement. In May 1999, however, the first international workshop was held to test how well the then current crystal structure prediction methods performed for four molecular compounds. The workshop was organised by the Cambridge Crystallographic Data Centre.\(^6\)
While only partially successful and with the conclusion that “at the present stage of development, perhaps the best that can be expected from crystal structure prediction programs is to provide a list of possible candidates for experimentally observable polymorphs”[6] it is now clear that much had been achieved in the intervening years. The most recent “Blind Test” in October 2015[7] (with five target systems covering a range of scenarios from a hydrated salt to a bulky flexible molecule) was summarised by Gibney[8] as “Chemists have succeeded at a fiendish task – how complex molecules will assemble in 3D”.

The overview summary of the 2015 exercise reports “Significant progress has been seen in treating flexible molecules, usage of hierarchical approaches to ranking structures, the application of density-functional approximations, and the establishment of new workflows and “best practices” for performing CSP calculations”. The paper also drew attention to the key fact that the results also highlighted the real advantages of the approach in complementing and augmenting experimental work. In particular the application of density functional methods in the energy ranking of putative structures clearly providing an effective strategy to enable more accurate ranking of the hypothetical structures.[9–12] In starting this article with the Dunitz paper and the topic of crystal structure prediction I also want to highlight that it is important to recognise that crystal property prediction is equally vital when organic solids with designed properties are sought – a requirement of true crystal engineering[13,14] (in pharmaceuticals, pigments, agrochemicals, electronics etc.) – and when effective organic solids are to be developed with minimum experimental work.[15]

The present paper will be primarily focused from the standpoint of pharmaceutical materials science. It has as its premise the concept of “Molecules, Materials, Medicines” reflecting “the idea that drugs are formed from a convergence of synthetic chemistry, materials science and engineering coupled with pharmacological and clinical evaluation”.[16]

By touching on some of the key aspects behind the development of new medicines (either in terms of the final product or during the purification of an important synthetic intermediate) the objective is to examine some of the challenges which exist in the area of pharmaceutical materials science. The approach will be to examine key aspects of developing a marketable medicine from a newly discovered active drug molecule including topics such as: polymorphism; crystal structure prediction; methods of screening for different forms (e.g. polymorphs, hydrates, solvates, and other multicomponent crystals
(e.g. cocrystals)); the issue of thermodynamic versus kinetic control and nucleation
effects;\textsuperscript{[17]} particle engineering;\textsuperscript{[18][19]} surface effects\textsuperscript{[20]} and the likely correlation between
form stability (chemical and physical) and lattice imperfections.\textsuperscript{[21]}

2. Polymorphism and form diversity

Polymorphism is ubiquitous amongst pharmaceutically relevant molecules\textsuperscript{[22–24]}. A
recent extensive review by Cruz-Cabeza et al\textsuperscript{[25]} has examined this issue in detail by using
the Cambridge Structural Database (CSD) as well as in-house pharmaceutical company
databases (at Hoffmann-La Roche and Eli Lilly and Company). In addition they undertook a
comparative study of 446 polymorphic crystals with energies and properties determined
with DFT-d methods. The situation is complex if not only because the CSD is likely to be
biased against the recording of polymorphs in the sense that historically a crystal structure
was determined only once in order to obtain molecular configuration and geometry
(particularly true for synthetically-challenging complex molecules), with no incentive (or
ability in terms of availability of the molecule) to screen for polymorphs.

Cruz-Cabeza et al, however, were able to report some firm conclusions. Molecular
flexibility or size (an area where the information from the industry databases was
important) did not correlate with likelihood of polymorphs being reported). For single-
component compounds the occurrence was 37 ± 12% in the CSD compared with 53 ± 12% in
the Roche dataset and 66 ± 11% in the Lilly dataset. This, they suggest, results from the
more extensive polymorph screens undertaken by industry as well as the reliance on other
methods of detecting polymorphs (e.g. powder X-ray diffraction and FTIR) clearly not
applicable to the CSD entries. Chiral molecules, on the other hand, were shown to be less
prone to polymorphism than their achiral counterparts and the existence of hydrogen
bonding functionality only marginally affected the extent of polymorphism. Bearing in mind
the significant impact that the unexpected appearance of a new polymorph may have on
the delivery of a drug molecule\textsuperscript{[26]} there is much to be learnt and digested from this
excellent study.

Cruz-Cabeza and Bernstein have also re-examined the situation of conformational
polymorphs.\textsuperscript{[27]} They provide an unambiguous definition of conformational change and
conformational polymorphism and the important requirement that there must be an energy
barrier to be overcome in order to claim two distinct conformers – in other words slight
adjustment of conformation by lattice forces whilst the molecule remains in the same potential well would not be sufficient – as they state “conformational change” as opposed to “conformational adjustment”.

From an experimental viewpoint it remains unclear, however, how the nature of the solvent influences conformational polymorphism. In a recent example of the case of tolfenamic acid, which experiences rapid fluctuations between conformers, a detailed experimental and computational study showed no clear link between solution chemistry and polymorph outcome. [28] The paper highlights a key challenge that exists for this system in linking the effects of solvent on conformation population in solution and the subsequent nucleation and growth of a conformational polymorph. [29] In the case of large drug-like molecules with significant conformational flexibility this remains an important challenge.

In reality little can be said about the extent of likely polymorphism for a particular molecule in the absence of an extensive experimental polymorph screen, [30] and even then with certainty that all forms have been isolated. [31] Screening can be by high-throughput methods [32,33] and/or additionally by more time-intensive approaches such as liquid and polymer assisted grinding, [34–37] melt growth, [38] interfacial crystallisation, [39] polymer heteronuclei, [40] epitaxial growth on substrate surfaces. [41] Additionally there will remain the intriguing issue of “disappearing polymorphs” and the possible presence (or absence) of adventitious (or “invisible”) seeds. An issue initially addressed by Dunitz and Bernstein [42] and recently further discussed by Bucar et al, who review both the scientific literature and patent litigation examples. [43] They draw attention to recent examples of polymorphs which (i) had been previously obtained but unexpectedly impossible to re-obtain and (ii) an initial inability to obtain crystal forms for which there was no a priori reason not to have expected them from a relatively straight-forward screen. Their sobering conclusion is that “it can never be stated that the most stable form has been found; at best it can be determined which of the known forms is the most stable” (my italics). When coupled with possible competing hydrate and solvate forms appearing during solution crystallization (and their different stoichiometry as well as polymorphs of the same stoichiometry) it is fair to conclude that this topic has still much to be explored. [44]

3. Crystal structure prediction
Dunitz concluded in 2003\footnote{1} “We have a long way to go before we can think of predicting which polymorph will be obtained under any given circumstances.” While it may be possible to predict alternate structures we cannot prescribe at the outset the approach necessary to obtain a particular structure. Nevertheless the range of putative structures generated might be studied to determine whether an observed polymorph was, within the limits of the accuracy of the calculations, the likely thermodynamically stable polymorph – assuming that the trend in lattice energy for different polymorphs is mirrored in free energy values.\footnote{7,45,46} Perhaps less crucial to the confidence in the absolute values of energy or density, the output might suggest that the observed structure is sufficiently lower in energy than other potential structures that there might be confidence in developing this form – provided its physical and chemical characteristic were adequate.

An alternate informative outcome might be that if the apparent lowest energy polymorph failed to be developable then it was unlikely that an alternate polymorph would be available without the risk (perhaps small, perhaps not) of it converting to the more stable form with time, moisture or subsequent handling. A third outcome might be the scenario where the observed polymorph was a member of a cluster of potential structures with similar energies. Would this mean that during scale-up concomitant polymorphism would be possible? The concept of “predictability” then might be best described in a more practical sense as “what are the realistic alternatives”.

To this end a step might be to examine, using various tools, other clues as to uncertainties concerning an observed polymorph. Here concepts such as molecular recognition\footnote{5} and synthon analysis\footnote{47} would be important.\footnote{48} Indeed, while CSP holds promise the increased levels of computational complexity associated with drug-like molecules possessing, for example, numerous degrees of molecular flexibility and increased likelihood of solvate formation, remains highly challenging.\footnote{49} Additionally for highly flexible molecules effects of solvent on relative conformer energies may itself alone be sufficient to emphasis kinetic rather than thermodynamic control. Less computationally demanding approaches for interrogating plausible structures include synthon analysis of predicted structures\footnote{47,50} and an examination of the “supermolecules” which might be generated\footnote{51} and which might pack effectively with respect to the individual components.\footnote{52} The hydrogen bond propensity (HBP) methodology,\footnote{53} which is based on a statistical comparison of the observed interactions of molecular fragments within an observed polymorph or
polymorphs against those seen for known structures within the crystal structure database is one approach. As demonstrated by Arlin et al locating a low energy polymorph from the CSP output and being able to study its predicted crystal arrangement may be used to develop a process for obtaining that polymorph through some crystallisation strategy circumventing the normal nucleation and growth process.\textsuperscript{41,54}

4. Multicomponent systems

While there are many contributing factors to the high attrition rate associated with drug development and production to a marketable drug product,\textsuperscript{55} low aqueous solubility is a frequent issue.\textsuperscript{56} As a consequence poor bioavailability is a common challenge requiring various approaches to increase the amount of drug substance absorbed into the blood stream. In the BCS classification\textsuperscript{57} this would typically be Class II drugs with low solubility and high permeability.

Amongst the general approaches utilised are the use of metastable polymorphs, salts, added surfactants, cyclodextrin, solid (polymer) dispersions and amorphous forms.\textsuperscript{58} One area where various inherent disadvantages are circumvented is in the area of multicomponent solids and in particular the recent interest in cocrystals \textsuperscript{59–62} and cocystal salts and ionic cocrystals\textsuperscript{63–65} A successful selection of a cocrystallizing molecule (coformer) from a very large list (based on GRAS (Generally Regarded as Safe) or EAFUS (everything added to foods in the United States) is also a challenge when only small amounts of material are available. Several approaches have been proposed,\textsuperscript{66} and utilised successfully to varying degrees, including the use of HBP calculations (described above) where self-self interactions are compared with possible drug-coformer interactions.\textsuperscript{67–70} Electrostatic potential energy mapping\textsuperscript{71} (aligned to the idea of the strongest hydrogen bond donors will ideally pair with the strongest hydrogen bond acceptor,\textsuperscript{72} and molecular similarity in terms of shape, size and polarity and other molecular identifiers.\textsuperscript{73}

In such systems there are three distinct issues. (i) Which potential pairs will actually cocrystallise rather than separate to individual crystals?\textsuperscript{74,75} (ii) What will be the preferred stoichiometry (for systems based on neutral molecules and unlike the case of salts there will be no required balancing of species)?\textsuperscript{76} (iii) Will relative lattice energy/density considerations require a comparison of the different polymorphic forms for each possible stoichiometric value?
CSP of such multicomponent systems would seem an attractive route but CSP becomes increasingly challenging when it involves more than one (potentially flexible) molecule in the asymmetric unit – either for a single component system with \( Z' = 2 \) or more or a multicomponent system (when \( Z' \) will be at least 2).\(^{45,77–80}\) The synthon approach\(^{47,50,52}\) can help in this regard in that a tentative “supermolecule” may be created between the two components (reducing the problem ideally again to a \( Z' = 1 \) situation)\(^{81}\) but as demonstrated by Bucar et al.\(^{82}\) using H/F interchange,\(^{83,84}\) this approach is less than ideal if there is a range of hydrogen bond donors and/or acceptors present.

5. Particle engineering, surfaces and defects

While the discussion so far has been based on internal structure, crystal size and habit are also important from a development viewpoint\(^{85}\) (e.g. ease of filtering and particle flow)\(^{86}\) but also in pulmonary drug delivery where particle size and size distribution is crucial\(^{87}\) and in the formation of nanosized crystals for direct targeting of active sites.\(^{19,88,89}\) Strategies include crystallization from supercritical fluids,\(^{90}\) milling, spray drying and sonocrystallization.\(^{91}\)

Coupled with this has been a development of analytical methods for characterisation of shape, surface chemistry and potential polymorph variation with decreasing particle size. TEM methods have provided examples where additional polymorphs have been detected,\(^{92}\) either as a result of its ability to obtain diffraction patterns from very small (sub-micron) individual crystals or possibly as a result of the fact that the crystal growth methods required to obtain suitably thin crystals (less than ca. 300nm) has resulted in the production of polymorphs not generated using growth methods aimed at larger crystals – frequently such large crystals require post-crystallisation milling to much smaller particles – a “top-down” approach”. Whether such a milling process (e.g. ball milling or jet-milling) will inadvertently result in polymorph conversion is likely to depend on the need for a solid-solid transformation pathway\(^{93}\) rather than perhaps a preferred nucleation process involved with the growth of small crystals.\(^{94}\) The top-down approach also produces high surface energy particles, and possibly highly defective particles and stress-induced phase transformations.\(^{95}\) A more ideal approach would be the “bottom-up” approach of obtaining directly small uniform particles e.g. by rapid precipitation from a solution\(^{96}\) using anti-solvent or rapid de-solvation e.g. spray drying.\(^{97}\) The use of tailor-made additives
(auxiliaries)\textsuperscript{[98]} is also a known way of modifying habit and particle size but may be limited in the case of pharmaceuticals due to final drug-substance purity issues.\textsuperscript{[99]} There will only be a finite (and probably small) number of potential impurity molecules which could be added to the final drug form.

An understanding of nucleation phenomenon remains a changing area of investigation. Experimental data frequently fails to follow kinetics modelled on classical nucleation theory and a two-step model (initially proposed for protein crystallisation) would seem to be more appropriate.\textsuperscript{[17,100]}

The need to fully understand surface effects is also important, especially from the viewpoint of tablets and inhalation devices where interfacial interactions are important.\textsuperscript{[101]} The enhanced instability of a drug crystal in contact with other excipients in a tablet is well known.\textsuperscript{[102]} The balance of cohesive and adhesive energy in inhalation formulations is also important. It is clear that knowledge of bulk behaviour (measuring essentially the internal properties of the crystal) can be very different to changes occurring at the external surface.\textsuperscript{[20,103]} Understanding growth features during crystallisation is also important. In this regard atomic force microscopy (and variants) is an important analytical tool\textsuperscript{[104–109]} – it continues to demonstrate interesting variation between bulk and surface reactivity as well as important information concerning crystal growth and dissolution.\textsuperscript{[110]}

Early studies recognised the potential role of imperfections in modifying reaction pathways in organic crystals, at the time being especially focused on rationalising apparent non-topochemical reactions.\textsuperscript{[111]} While the presence of defects (or simply structural disorder) in pharmaceutical materials is recognised\textsuperscript{[112]} it remains uncertain as to their nature and extent.\textsuperscript{[113]} Defects are frequently referred to in a general way without clear identification of the exact local structure created by their presence. Classical examples include dislocations (both edge and screw) with concepts based on our understanding of such defects in metals. But the local structure around such faults is likely to be highly complex,\textsuperscript{[114]} with conformational flexibility being one variable not encountered in simple metals. Even the simple case of a vacancy is likely to be significantly more complex given the possibly of local reorganisation. While line defects and planar defects (e.g. twin and domain boundaries) may be visualised using transmission electron microscopy or X-ray topography,\textsuperscript{[21]} less well defined defect regions will be extremely hard to visualise or indeed
model but are likely to be nucleation centres for subsequent physical conversion or chemical reactivity.

6. Concluding remarks and the future challenges

Developing a newly discovered active molecule to a marketable drug product is associated with a long development time and a high attrition rate. This can be a result of pre-clinical toxicology issues or early clinical failures.[115] It is also the case that formulators are able to modify drug product characteristics to improve performance issues.[116,117] It remains important, however, for solid formulations that the associated solid state chemistry is sufficiently understood that the development of scale-up processes, production and long-term manufacturing are as robust as possible. The emergence of new polymorphs and the crystallisation of amorphous forms are just two examples of where major issues can arise.[26,118] A better understanding of solid state phase transformations would enable more confidence in developing metastable polymorphs, stabilizing amorphous forms[119] and further work in identifying the types and role of defects in modifying degradation pathways will also be important.[120] Certain crystal attributes can be modelled reasonably well – habit,[121] mechanical properties,[122] dislocation nucleation and dynamics[123,124] - and importance in modelling and predicting solubility.[125][126] Combining crystal structure with associated property predictions is likely to be a key driver for future research efforts towards establishing robust development programs.

7 Acknowledgments

I am grateful to Dr Dritan Hasa for assistance in preparing this article.

8 References


