European Journal of Pharmaceutics and Biopharmaceutics 71 (2009) 23–37

Contents lists available at ScienceDirect

European Journal of Pharmaceutics and Biopharmaceutics

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

Review article

Solid form screening – A review

Jaakko Aaltonen [a,*], Morten Allesø [b], Sabiruddin Mirza [c], Vishal Koradia [b], Keith C. Gordon [d], Jukka Rantanen [b]

[a] School of Pharmacy, University of Otago, Dunedin, New Zealand
[b] Department of Pharmaceutics and Analytical Chemistry, University of Copenhagen, Denmark
[c] Division of Pharmaceutical Technology, University of Helsinki, Finland
[d] Department of Chemistry, University of Otago, Dunedin, New Zealand

ARTICLE INFO

Article history:
Received 29 February 2008
Accepted in revised form 18 July 2008
Available online 31 July 2008

Keywords:
Amorphous
Hydrate
Polymorph
Polymorphism
Screening
Solid form
Solid state
Solvate

ABSTRACT

Solid form screening, the activity of generating and analysing different solid forms of an active pharmaceutical ingredient (API), has become an essential part of drug development. The multi-step screening process needs to be designed, performed and evaluated carefully, since the decisions made based on the screening may have consequences on the whole lifecycle of a pharmaceutical product. The selection of the form for development is made after solid form screening. The selection criteria include not only pharmaceutically relevant properties, such as therapeutic efficacy and processing characteristics, but also intellectual property (IP) issues. In this paper, basic principles of solid form screening are reviewed, including the methods used in experimental screening (generation, characterisation and analysis of solid forms, data mining tools, and high-throughput screening technologies) as well as basics of computational methods. Differences between solid form screening strategies of branded and generic pharmaceutical manufacturers are also discussed.

1. Introduction and background

In a classic paper by Haleblian and McCrone, polymorphism was defined as the ability of any compound to crystallise as more than one distinct crystal species [1]. However, in physical pharmacy the word 'polymorphism' is nowadays often used to cover a variety of solid forms of active pharmaceutical ingredients (APIs) and excipients including crystalline, amorphous, and also solvate/hydrate forms. Accordingly, the activity of generating, isolating and analysing different solid forms of an API is known as polymorph screening. In this paper, to avoid term confusion, all above-mentioned solid modifications are referred to as solid forms. The aim is to give a pharmaceutical outlook on solid form screening – what it consists of and what is the significance of it for the pharmaceutical field. As such, focus is not put on the screening and selection of salts and co-crystals.

1.1. Solid form screening

Nowadays solid form screening is a standard procedure in drug development, but it was not until the last decades of the 20th century that the whole pharmaceutical industry became aware of polymorphism even though the phenomenon had been known since the early 19th century [2]. The patent cases [3] and enormously expensive Hatch–Waxman agreements between branded and generic drug manufacturers [4], and production problems such as the sudden appearance of another form of ritonavir (with low solubility and bioavailability) [5] were needed for the industry to start taking polymorphism seriously. Regardless of the fact that solid form screening is a regulatory requirement for new pharmaceuticals [6], the whole topic remains still somewhat open since no universal guidelines on solid form screening can be written because every compound possesses unique properties. Further, all polymorphism-related phenomena, e.g., nucleation, are not yet fully understood [7,8].

The aim of solid form screening is to find the optimal form with the best characteristics for development. In order to be able to select the optimal form, knowledge of as many as possible forms is needed. The choice of the form for development is usually a compromise between physical, chemical, pharmaceutical and biopharmaceutical properties (see Section 1.2). Traditionally, the most stable form is favoured over other forms because of its lower tendency to solid phase transformations. It is important to identify the stable form as early as possible in the drug development process to avoid subsequent setbacks [9,10]. However, metastable forms are sometimes deliberately chosen – usually for better solubility and thus bioavailability [11]. Solid form screening should not be a one-time effort performed only during the preformulation stage

* Corresponding author. School of Pharmacy, University of Otago, P.O. Box 913, Dunedin 9054, New Zealand. Tel.: +64 3 479 7272; fax: +64 3 479 7034.
E-mail address: jaakko.aaltonen@otago.ac.nz (J. Aaltonen).

0939-6411/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejpb.2008.07.014

of drug development. Instead, the solid state of the API should be monitored in a continuous fashion also during scale-up and manufacturing, to detect the occasional event of late appearing polymorphs as early as possible.

Solid form screening can be approached experimentally and computationally. Experimental screening consists of the preparation step, during which various forms are generated and isolated (see Section 2.1), and the analysis step, which involves the use of various measurement techniques (see Section 2.2) and also the data analysis (see Section 2.3). Computational methods of polymorph prediction (see Section 3) have evolved markedly during the last years, but one cannot fully rely on them yet. Even if the crystal structures are not predictable, computational methods can be used to help rationalise the experimental procedures [12] and decide whether the stable form has been found or not [13]. Solid form screening is one of the cornerstones of drug development and plays an important part in pharmaceutical product lifecycle management [14]. If most of the relevant forms are found the innovator company can achieve a very valuable intellectual property (IP) situation. Poorly conducted screens and unsuccessful patenting strategies on the other hand open new possibilities for competitors (see Section 4).

1.2. Differences between solid forms and their implications on physical, chemical, pharmaceutical and biopharmaceutical properties

In the crystalline state (polymorphs, solvates/hydrates, co-crystals), the constituent molecules are arranged into a fixed repeating array built of unit cells, which is known as lattice, whereas in the amorphous state there is no definite long-range order [15,16]. Polymorphs contain molecules of only one chemical species in their unit cells. There are two ways in which crystal lattices can form: packing polymorphism and conformational polymorphism. If rigid molecules with a specific conformation are packed in different arrangements, then it is referred to as packing polymorphism, whereas conformational polymorphism denotes crystal forms that consist of flexible molecules with different conformations packed in different arrangements [16]. See paracetamol [17] and L-glutamic acid [18] for examples of packing and conformational polymorphism, respectively.

In case the unit cell is built of host molecules (the API) accompanied by guest molecules the resulting solid form is called a solvate or a co-crystal depending on whether the guest species is liquid or solid at ambient conditions, respectively [19]. If the guest molecule is water then the term hydrate is used. (There has been, and still is, controversy in the literature regarding the nomenclature of multi-component crystals.) Solvates and co-crystals can also exhibit polymorphism. Hydrates are the most common solvates and deserve special attention. Many APIs are capable of forming hydrates due to the small size and multidirectional hydrogen bonding capability of the water molecule [20,21]. Hydrate formation stabilises the crystal structure via intermolecular bonding in the crystal lattice, resulting in hydrates being the stable forms in aqueous surroundings (below dehydration temperature). Many hydrates are also stable in ambient conditions, and therefore are quite often chosen as the form for development. The hydrate form can be chosen for development to avoid hydrate formation during downstream processing, but then again with hydrates there is a risk of dehydration. A classification system of hydrates has been suggested [22].

Amorphous state can be the result of either physical manipulation or the intrinsic nature of the compound [23]. The amorphous form is the most soluble form, but on the other hand, it is also the form with the lowest stability. The current trend of new APIs getting less and less soluble has given rise to the use of metastable forms in formulations and also another research topic, stabilisation of metastable forms and formulations thereof [18,24,25]. The physical stability issues are the main hurdles in developing formulations of amorphous APIs [26]. Regardless of being stable in "dry" conditions over the whole shelf-life, metastable forms may rapidly convert to the stable form upon administration, since solvent-mediated solid phase transformation kinetics are faster than those of solid-solid transformations. Therefore, the solubility advantage of amorphous (and other metastable) forms may not always be fully exploited [27].

Table 1
Physical properties that differ among various solid forms

Packing properties	Unit cell volume (crystalline forms only), density, refractive index, hygroscopicity
Thermodynamic properties	Melting point, enthalpy, entropy, free energy, solubility
Spectroscopic properties	Electronic transitions (UV–vis spectra), vibrational transitions (IR and Raman spectra), rotational transitions (far-IR or microwave spectra), nuclear spin transitions (NMR spectra)
Kinetic properties	Dissolution rate, rates of solid-state reactions, stability
Surface properties	Surface free energy, interfacial tensions, crystal habit
Mechanical properties	Hardness, tensile strength, compactibility, tableting, flowability

Modified from [16].

The solid forms of a given API can have significantly different physicochemical properties that can affect its performance [1,28]. Some of these properties are listed in Table 1. If solubility and/or dissolution rate are dependent on the solid form, the bioavailability of the API can be affected. This is a particularly important note when developing BCS class II APIs (low solubility and high permeability) with dissolution dependent bioavailability [29]. Examples of APIs with bioavailability problems due to solid-state phenomena are carbamazepine [30,31] and ritonavir [5]. Mechanical property differences can affect processing behaviour, and this is the case, for example, with paracetamol [32]: direct compression of form II is feasible, whereas with form I, binder excipients have to be used [33]. Also different forms of theophylline [34] and sulfamerazine [35] have been reported to show different processing characteristics. Stability is a very important property of a solid form, considering that raw materials and pharmaceutical products may be stored for prolonged periods and the solid state must remain unchanged. In addition to physical stability, chemical stability also has to be taken into account. Chemical reactivity can vary between different solid forms, and sometimes certain solid forms can be used to cause reactions when desired or to prevent reactions when they are to be avoided [36]. For examples of differences in chemical stability between forms see prednisolone *tert*-butylacetate [37] and quinapril HCl [38].

2. Experimental solid form screening

2.1. Generation of solid forms

2.1.1. Basics – thermodynamics and kinetic effects

Crystallisation is the key experimental technique used to execute solid form screens. Within this review, crystallisation is considered as a tool to generate multiple solid forms of a pharmaceutical compound; the fundamental theoretical aspects of this process can be found elsewhere [39,40]. Classically, the crystallisation process is described in terms of two distinct steps, nucleation and crystal growth [41], with the resulting physical form being the consequence of the kinetic relationship between these two elementary processes. In other words, for a polymorphic system the polymorph that nucleates first is thought to come from the cluster that exhibits the fastest nucleation rate as a result of its lowest free energy barrier (ΔG) to nucleation. However, the nature

of the polymorph that eventually crystallises is determined by the combination of the relative nucleation rates and the relative crystal growth rates of the polymorphs [42,43]. Hence, nuclei of different structures can form and coexist in a given crystallisation leading to a mixture of crystalline forms in the resulting solid when kinetic factors prevent the achievement of equilibrium. It is therefore a general rule that the stable form is likely to be obtained when operating under thermodynamic conditions (e.g., slow cooling), while a metastable form is expected to be produced under kinetic conditions (e.g., rapid cooling). In the context of polymorph screening, this implies that the interplay between kinetic and thermodynamic factors has to be extensively employed to discover all the relevant solid forms of a compound in question. The possible scenarios of crystallisation in a dimorphic system are schematically presented in Fig. 1.

2.1.2. Different ways to generate solid material

There are a large number of classical techniques that can be used to generate solid materials [25,40,44,45] (Table 2). Since one approach can favour the nucleation and growth of one form over another, it is essential to perform screening experiments by a variety of methods under various process conditions. In general, crystallisation from solution (cooling or evaporative, as well as slurry conversion) and recrystallisation from a neat compound (sublimation, thermal treatment, crystallisation from the melt, grinding, and thermal desolvation) are the methods of choice for solid form screens [45,46]. Crystallisation from solution is customarily used in solid form screens for several reasons. Firstly, a large number of polymorphs and solvates can be discovered by changing the solvent system. Secondly, pharmaceutical solids are often exposed to different solvents during processing, and thus the solvate/hydrate formation tendency of APIs should be systematically studied in order to design the manufacturing processes. Thirdly, desolvation of solvates or dehydration of hydrates is another useful, and sometimes the only, technique to discover a polymorphic form [47,48]. Finally, a solvate (and especially a hydrate) can be of interest as a commercial product.

Table 2
Methods to generate various solid forms [25,40,44,45]

Method	Degrees of freedom
Crystallisation by cooling a solution	Solvent, cooling profile, concentration, mixing
Solvent evaporation	Solvent, initial concentration, evaporation rate, temperature, pressure, ambient relative humidity
Precipitation	Solvent, anti-solvent, rate of anti-solvent addition, mixing, temperature
Vapour diffusion	Solvents, temperature, concentration
Suspension equilibration	Solvent, temperature, solubility, temperature programs, mixing, equilibration time
Crystallisation from the melt	Temperature changes (min, max, gradients)
Quench cooling the melt	Cooling rate
Heat induced transformations	Temperature changes
Sublimation	Temperature gradient, pressure, surface type
Desolvation of solvates	Temperature, pressure
pH change	Temperature, rate of change, acid/conjugate base ratio
Mechanical treatment (i.e., milling, cryo-grinding)	Milling time, mill type
Freeze-drying	Solvent, concentration, temperature programs
Spray drying	Solvent, concentration, drying temperature

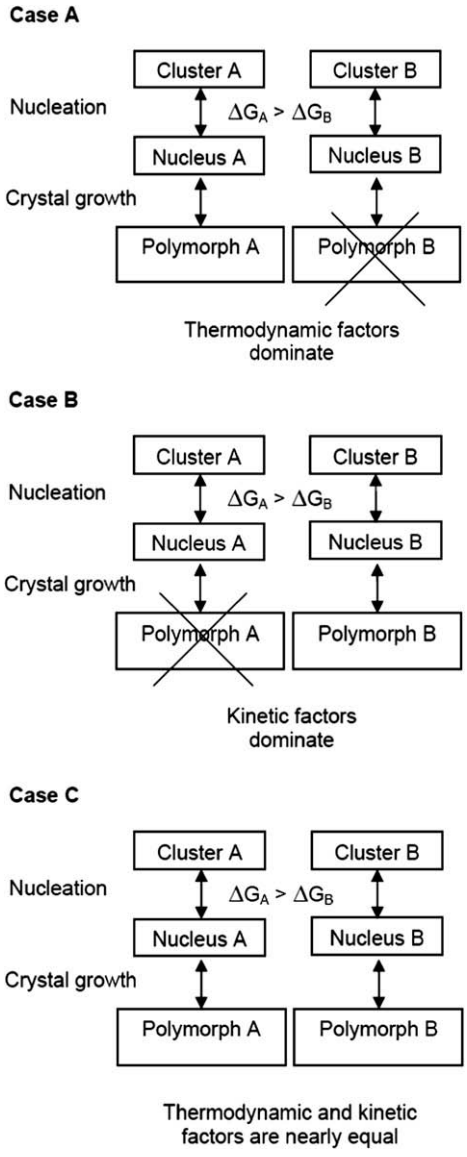

Fig. 1. Possible scenarios of crystallisation in a dimorphic system under different conditions. Polymorph A is the stable form with a higher free energy barrier (ΔG) to nucleation.

Recently, several innovative techniques – for example, capillary crystallisation [49], laser-induced crystallisation [50], and sonocrystallisation [51] – that promote nucleation and hence discovery of alternate crystalline forms have been reported. Another recent solid-state topic of growing interest is the design and formation of co-crystals. Co-crystallisation opens new ways to produce solid forms with unique characteristics for, e.g., combination therapy and IP protection. For deeper insight into the preparation of co-crystals the reader is referred to texts by Rodríguez-Hornedo et al. [52] and Stahly [53].

2.1.3. Critical parameters in crystallisation of polymorphs

As crystallisation is a highly complex process there are several process variables that can affect the outcome (Table 3). Supersaturation as the driving force of crystallisation is the key thermodynamic variable that affects the kinetics of crystal nucleation and growth. In other words, the resulting crystalline form may vary with the degree of supersaturation. Temperature can be considered as the second most significant variable affecting crystallisation outcome in a polymorphic system. The effect of temperature has both thermodynamic and kinetic implications, particularly for enantiotropic polymorphs which change the solubility order near the transition temperature [54]. Practically speaking, crystallisation at one temperature may produce one polymorph, while crystallisation at another temperature may yield a second polymorph [55]. For some solvate systems,

Table 3
Process variables affecting the outcome of crystallisation from solution [46]

Crystallising phases		Crystallisation method			
Polymorphs/solvates	Salts/co-crystals	Cooling crystallisation	Evaporation	Precipitation	Slurry conversion
Degree of supersaturation	Counter-ion type	Heating rate	Evaporation rate	Anti-solvent type	Solvent type
Solvent composition	Acid/base ratio	Cooling rate	Evaporation time	Rate of anti-solvent addition	Incubation temperature
Additive type	Solvent/Solvent combination	Maximum temperature	Carrier gas	Temperature of anti-solvent addition	Incubation time
Additive concentration	Degree of supersaturation	Incubation time	Surface-volume ratio	Time of anti-solvent addition	Thermal cycling and gradients
	Additive type Additive concentration pH Ionic strength	Incubation temperature			

it can be anticipated that by changing the harvesting temperature the solvates of various stoichiometry will be obtained, with lower temperatures favouring the formation of solvates of higher stoichiometry [20]. Therefore, crystallisations under various temperature profiles should be performed during solid form screens. Solvent, additives (and impurities), interface, pH and host-guest composition have been classified as secondary factors affecting the crystallisation outcome [56], mainly through their effect on the degree of supersaturation (Fig. 2).

Selecting the right set of solvents is important when designing solid form screening experiments. Selection criteria should encompass (1) solvents with broad distribution of properties [57–59], (2) potential solvents used in synthesis, purification and processing, (3) solvents and excipients used in the final formulation. (E.g., if the pharmaceutical product is formulated in a soft gelatin capsule, the carrier medium should be included in the screening.) In addition, crystallisations from water and water-solvent mixtures are usually included in the screens to generate hydrates. Papers dealing with the analysis of multivariate solvent databases have been published to aid the selection of solvents [57,59,60].

In summary, to increase the probability of discovering all the relevant forms, the multiparameter space that contributes to solid form diversity should be covered as broadly as possible. This is usually achieved by designing a rational set of the process variables [61].

2.1.4. Seeded and additive-mediated crystallisations

Seeding (the addition of solid particles of the desired phase to a crystallisation medium) is a common approach to crystallise the desired crystalline form [62]. Controlling crystallisation outcomes via seeding relies on the potential of crystal surfaces to promote heterogeneous or secondary nucleation, while avoiding heterogeneous nucleation mediated by unknown contaminants [40]. In the context of solid form screening, this implies that some precaution measures (e.g., filtration of a supersaturated solution) should be undertaken to avoid the crystallisation mediated by the initial solid phase. Conversely, it may be useful to employ heterogeneous nucleation [18,24,63,64] mediated by various additives (e.g., structurally related compounds or polymers) as a solid form screening tool. In such cases, the additive acts as an additional diversity element which affects the crystallisation outcome, and thus provides the potential for the discovery of unknown solid forms without prior knowledge of the crystal structure. For example, a fourth polymorph of carbamazepine, which appeared to be more stable than the well-studied trigonal form [65], was discovered using the polymer heteronucleation approach [66].

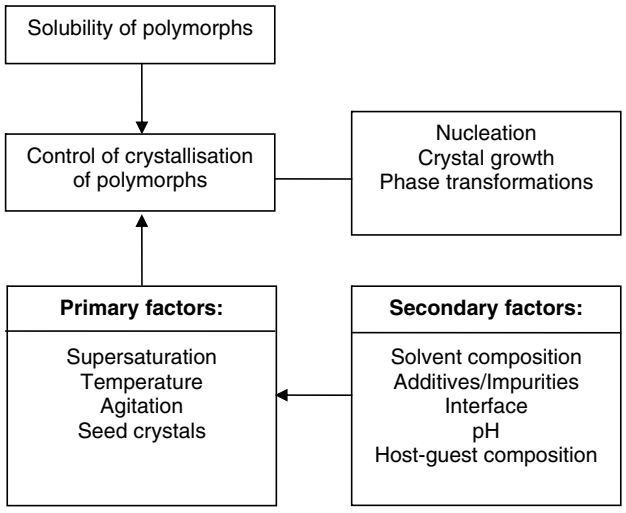

Fig. 2. Hierarchy of the process parameters controlling crystallisation outcome in polymorphic systems. Modified from Kitamura et al. [56].

2.1.5. High-throughput screening technologies

High-throughput (HT) screening utilises fully automated robotic systems capable of performing thousands of crystallisations per week with only few grams of API consumption [44,67]. It is believed that the immense number of crystallisation trials will increase the probability of finding solid forms suitable for further development and/or patenting. HT screening is most commonly carried out in 96-well plate systems in which the particular solid, dissolved in a suitable solvent, is initially dispensed using automated liquid handling systems. The amount of API is reported to lie in the range of 0.5–10 mg per sample [67–69]. Different levels of supersaturation can be achieved by, for example, heating/cooling, evaporation, and by varying the nominal concentration of API. Similarly to bench-scale crystallisations, slurry experiments can also be implemented in HT systems. The resulting solid is commonly analysed *in-situ* using either XRPD or Raman spectroscopy. The generated forms are thereafter identified using one or several suitable data mining tools (see Section 2.3) and the results are stored in a database for later use. It should be noted that microscale *in-situ* analysis is relatively prone to sampling errors, and may thus lead to bad quality data or false results. Very often follow-up studies (e.g., including up-scaled API quantities) are performed for the purpose of further elucidating interesting solid forms found in the initial HT screening or in an attempt to find new forms by further varying the crystallisation conditions. As an example of this, Peterson et al. demonstrated an approach to the so-called iterative HT screening comprising a multitude of experiments under varying as well as similar conditions [70]. Acetaminophen (paracetamal) was subjected to screening and three experimental runs/iterations of evaporative crystallisations were required to identify and further test the reproducibility of Form II formation (starting material was Form I). It was only possible to generate Form III by subsequent melt crystallisation. Conditions

and results for a series of HT polymorphism studies are summarised in Table 4. It is important to emphasise that in spite of the vast number of crystallisations, the hit rate (i.e., the percentage of vials containing API precipitates) is often very low. Transform Pharmaceuticals have reported hit rates between 2.5% and 13% for HT screening of various APIs [69–71]. Clearly, the ultimate advantage of HT solid form screening lies in the large number of experiments carried out, which in addition utilise low API quantities and little if no manual intervention. Automated approaches do, however, suffer from some drawbacks as opposed to manual bench-scale crystallisations. Firstly, there is limited ability to incorporate some of the pivotal crystallisation methods other than the basic solvent-based techniques (see Table 2). Secondly, a large fraction of the found solid forms are quite often not true polymorphs but rather solvates, and therefore less useful for further development (although this information may still be beneficial from a manufacturing point of view) [72]. Lastly, subsequent up-scaling of solvent volumes, as a part of further stability studies or API manufacturing, can change the polymorphic outcome markedly.

It is generally impossible to guarantee that all solid forms of a compound in question will be discovered by HT screening since there is no method that provides complete exploration of the solid-state landscape. It is therefore recommended to supplement the results of the HT studies with more in-depth crystallisation experiments.

2.2. Identification and analysis of solid forms

There is a wide array of methods available for solid phase analysis of pharmaceutical materials, and some excellent books and book chapters on the methods have been published [73–75]. Techniques commonly used to study solid-state properties are listed in Table 5 [76]. The method of choice for a specific case depends on the key parameters one needs to determine and how deeply they have to be investigated. Usually it is advisable to use two or more complementary methods to obtain a reliable knowledge of the forms, but as mentioned in the previous chapter the amount of samples produced by HT screening methods can be very large

Table 4
Conditions and results for a series of HT screening studies

Reference	API	Platform/company	No. of crystallisations/ no. of solvents	HT crystallisation method	HT analytical method	No. of forms found
Park et al. [67]	Buspirone hydrochloride	SSCI, Inc.	288 crystallisations/ solvents N/A	N/A	XRPD, Raman	3 polymorphs 3 solvates Amorphous phase
Florence et al. [72]	Carbamazepine	Chemspeed Accelerator SLT100	594 crystallisations/66 solvents	Cooling, evaporation, cooling and evaporation	XRPD	3 polymorphs (Forms I, II and III) 9 solvates
Desrosiers [68]	Cimetidine	Symyx technologies, Inc.	288 crystallisations/20 solvents, 84 solvent mixtures	Cooling, evaporation, precipitation	Raman, XRPD, polarising light microscopy	3 polymorphs
Hilfiker et al. [89]	Carbamazepine	Solvias AG	Crystallisations N/A/43 solvents and mixtures	Evaporation, suspension, desolvation	Raman	4 known polymorphs 2 known solvates (dihydrate and acetone) 5 suggested new forms (incl. 2 solvates)
Almarsson et al. [71]	Angiotensin II receptor antagonist MK-996	CrystalMax™/TransForm Pharmaceuticals, Inc.	1440 crystallisations/ 21 solvents and mixtures	Cooling (three nominal concentration levels), separate hydrate screen	Raman, XRPD	18 crystalline forms
Almarsson et al. [71]	Sertraline HCl	CrystalMax™/TransForm Pharmaceuticals, Inc.	3072 crystallisations/ 24 solvents and mixtures	Cooling (two nominal concentration levels)	Raman, XRPD	10 forms (4 additional forms by follow up studies)
Morissette et al. [69]	Ritonavir	CrystalMax™/TransForm Pharmaceuticals, Inc.	2000+ crystallisations/ 24 solvents (single and binary mixtures)	Cooling (four nominal concentration levels)	Raman	3 polymorphs (I, II, IV) 1 solvate (III). Follow up studies revealed a trihydrate (Form V)
Peterson et al. [70]	Acetaminophen	CrystalMax™/TransForm Pharmaceuticals, Inc.	7776 crystallisations (first iteration)/16 solvents (single and binary mixtures)	Cooling (three nominal concentration levels)	Raman	2 polymorphs (I and II). Form III was found in subsequent studies

Table 5
Methods to study solid-state properties

Method	Data measured	Property measured/use
X-ray diffraction (single crystal and XRPD)	Diffractogram	Crystallographic properties
Infrared (IR) spectroscopy	IR spectrum	Chemical information
Raman spectroscopy	Raman spectrum	Chemical information (complementary to IR)
Terahertz pulsed spectroscopy (TPS)	Terahertz pulsed spectrum	Chemical information, lattice phonon modes
Near-infrared spectroscopy (NIR)	Near-infrared spectrum	Chemical information (overtones and combinations of IR vibrations)
Solid-state NMR	Magnetic resonance	Chemical information
Differential scanning calorimetry (DSC)	Heat flow *vs.* temperature	Thermal events
Thermogravimetry (TG)	Change of mass *vs.* temperature	Solvate/hydrate studies
Microscopy, PLM, SEM	Microscopy under the influence of light or electron radiation	Morphology, surface examination, dehydration, polymorphism (PLM)
Moisture sorption/desorption isotherms	Change of mass *vs.* variable RH%	Hygroscopicity behaviour (hydrate formation, dehydration, amorphous crystallisation)
Solubility/dissolution	Amount dissolved in different solvents or temperatures/*vs.* time	Solubility/dissolution rate measurement
Microcalorimetry	Heat flow *vs.* time	Quantification of amorphous form
Solution calorimetry	Heat flow during dissolution	Quantification of polymorphs and amorphous form

and the use of multiple methods of analysis may not be reasonable timescale wise. Instead of using different methods to complement each other, it may be wise to use hyphenated techniques that permit complementary data to be acquired in a single measurement.

In addition to the identification and quantification of the solid forms, the hygroscopicity and moisture sorption behaviour as well as solubility and dissolution rate of the solid forms need to be tested to be able to select the optimal form. Moisture sorption analysers [77,78] and dissolution apparatuses [79] can be interfaced with *in-situ* spectroscopy to achieve better understanding of the solid-state phenomena during the analysis. The intrinsic dissolution rate (IDR) test [80] is a good way to investigate the effects of formulation excipients and gastrointestinal fluids (by the use of biorelevant dissolution media) on various solid forms and possible solid-state transformations occurring during dissolution. Generally speaking, it is wise to investigate the effect of excipients on the solid-state stability in both "dry" and "wet" conditions.

2.3. Data mining in solid form screening

An inevitable consequence of solid form screening is the generation of either small or large amounts of data. Effectively extracting relevant information from this data pool is no easy task for the researcher, considering the myriad of existing data analysis tools, ranging from simple visual inspection to unsupervised or supervised multivariate analysis. It is not possible to identify one gold standard for data analysis in solid form screening; instead the specific choice of approach depends on (1) the features of the data matrix with respect to matrix size and presence/absence of independent variables, (2) instrumentation used for data acquisition and (3) user preferences. In general, data treatment may be divided into the following categories: design of experiments, visual inspection of data, unsupervised and supervised analysis (Fig. 3). It is important that the researcher understands the possibilities and limitations of each category before using the data analytical tools.

2.3.1. Design of experiments

Before initiating any experiment, careful considerations should be paid as to how it should be carried out, that is, the construction of a statistical design. Failure to do so may result in poor quality data with little, if no, possibility of performing efficient and meaningful data analysis. The goal of design of experiments (DoE) is to plan and conduct experiments that allow extraction of a maximum amount of information from the collected data in as few runs as possible. As part of the DoE, (expected) critical process parameters/factors and their settings/levels are identified [81]. Ultimately, this will permit assessment of how each factor and their interactions affect a certain response variable through suitable statistical analysis. As reviewed by Yu et al. [82], DoE may be particularly useful for monitoring and controlling complex crystallisation processes, involving several critical process variables such as heating/cooling profiles, solvent type, solvent evaporation rate profiles, and seeding. These factors are highly interrelated and will affect the solid-state outcome of the crystallisation process. Thus, DoE can help unravel the complex nature of these and clarify their impact on the nucleation and crystallisation of several crystal forms at the early stage of solid form screening. An identical setup (merely implemented at larger scale) may be applied at a later stage for the controlled manufacturing of a desired API crystal form [82]. Al-Zoubi et al. [83] implemented a full factorial design to elucidate the effects of cooling temperature and harvesting time, each at three levels, on the formation and quality of orthorhombic paracetamol (Form II polymorph) via solution-mediated transformation. Data analysis on the factors and their interactions was carried out using ANOVA and revealed significant effects of both crystallisation parameters on several response variables as well as significant interaction. In a related study, the effects of various parameters on the solution-mediated polymorphic transformation of buspirone hydrochloride were investigated [84]. This experiment utilised a 2^4 factorial design having pH, solvent composition, amount of co-solvent, and impurities (seeding) as factors and degree of interconversion of metastable form 2 to stable form 1 as a response variable. The data were subjected to response surface modelling to conveniently visualise the effects of the four factors. From here, pH and amount of co-solvent and their interaction were found to have the strongest effect on the interconversion to form 1. Higher-order interactions are very common in chemical processes [85]. Unfortunately, interpretation of more than two-way interactions is almost impossible using traditional multi-way ANOVA and more sophisticated modelling tools are needed. In this context, GEMANOVA (GEneralised Multiplicative ANOVA) is suggested as a possible alternative to ANOVA for analysing complex data effectively and for obtaining more interpretable solutions enabling the overview of the whole sampling region. So far, GEMANOVA has been used mainly within the food industry [86]. However, it is expected to provide distinct advantages for the elucidation of complex pharmaceutical processes, such as crystallisation, in which the researcher wishes to understand and visualise the effect of

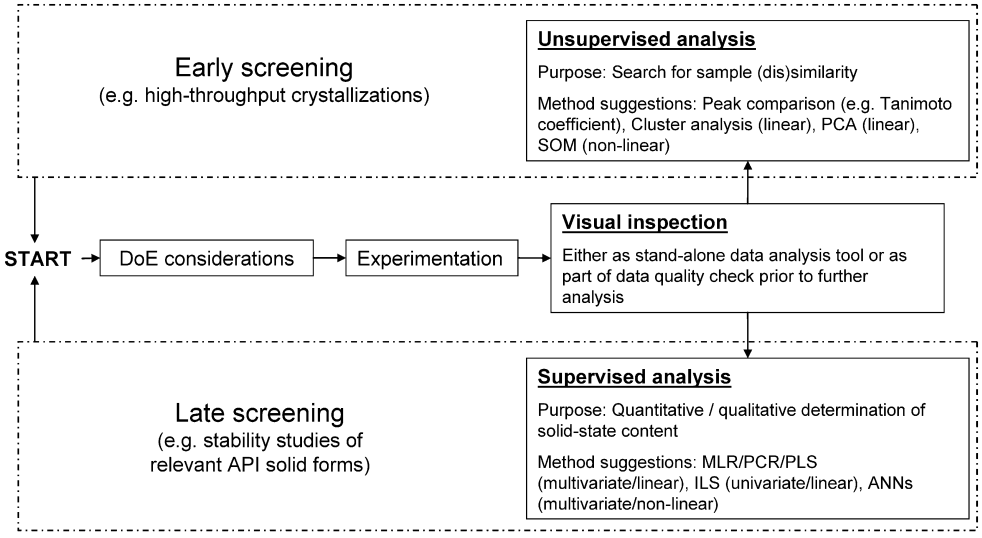

Fig. 3. Data analysis suggestions for solid form screening.

interacting process variables on, for instance, polymorphic outcome.

2.3.2. Visual inspection of data

Visual evaluation of experimental raw or pre-processed data is an often utilised approach in many areas of research and not least in the investigation of solid-state phenomena. Evidently, manual assessment seems most rational in cases where it is practically feasible, that is, when the amount of data is limited and/or if well-resolved characteristic peaks are present. It is important to point out that visual inspection of spectral data can sometimes be tedious, because the differences between the spectra of various solid forms can be very small. In cases where manual examination of data is unfeasible, it might be more rational to compare entire spectra by simple pair-wise subtraction, thereby permitting a more convenient overview of the spectral regions of interest for differentiating between the solid forms.

Visual inspection remains a powerful solution to the task of data interpretation. Even when applying sophisticated data modelling tools (see below), it is still highly recommended to support the findings by visual inspection of the relevant raw data. This greatly reduces the risk of drawing erroneous conclusions from the solid form screening experiments.

2.3.3. Unsupervised methods

Unsupervised analysis of data is performed when no *a priori* information on the solid material under investigation is available. This is often the situation during early screening of API candidates, where little is known about the solid-state properties [87]. In the current review, unsupervised methods are related to clustering of data by various mathematical equations/algorithms. The search for groupings or patterns in solid-state data is often initiated whenever the generated data matrix is too large for visual inspection. Evidently, the use of automated HT systems easily performing thousands of crystallisations requires effective data mining in order to provide a satisfactory overview of all the data in one single plot. During early screening, it is impossible to predict which specific regions of the diffractogram and/or spectrum (yet undiscovered) the solid forms will affect. Therefore, many of the unsupervised methods utilise more than just a single peak for the analysis. HT studies performed using the CrystalMax® screen have demonstrated a simple approach to unsupervised analysis of Raman data [69,70,88]. Here, spectral similarity was assessed by calculating the so-called Tanimoto coefficient for each pair-wise combination of Raman spectra. Tanimoto values close to 0 and 1 indicate high and low similarity, respectively, and the results are easily visualised using a contour-like plot. In general, the use of spectral peak positions, for the calculation of some correlation measure, is a widely accepted approach among researchers involved in HT screening [89]. This approach has also proven useful for treatment of XRPD data [71]. It is, however, recommended to show extra care when analysing XRPD data, since changes in crystallinity and preferred orientation can obscure the XRPD pattern markedly [90]. In these cases, careful filtering of the XRPD data is imperative to ensure successful data analysis [71].

For treatment of information-rich data, it can be advantageous to apply multivariate techniques. In multivariate analysis, the measured data are first organised into a data matrix, having samples/spectra as rows, and *x*-variables (wavelength/number, 2θ angle, etc.) as columns. Each combination of sample and *x*-variable will have an associated measured value (absorbance, intensity, etc.). Subsequently, this *X*-matrix is subjected to various multivariate techniques, utilising either the whole spectral range (all *x*-variables) or selected regions. In this respect, cluster analysis [91] has proven very useful in the search for data clusters in early solid form screening. This algorithm searches for clusters based on a predefined measure of (dis)similarity, typically a distance measure (e.g., Euclidean, Mahalanobis, city-block) or correlation coefficients. Sample (dis)similarity is hierarchically visualised in a dendrogram. Cluster analysis, with correlation coefficient as (dis)similarity) measure, has previously been used for analysis of XRPD [92] and Raman data [87] in HT screening experiments and for NIR data acquired from a bench-scale polymorph screening of sulfathiazole [93]. An associated problem with cluster analysis is to determine where to 'cut' the dendrogram, that is, to estimate the numbers of actual clusters [59,87]. Therefore, it is highly recommended to supplement the cluster analysis with other data analysis tools to ensure the correct interpretation of the experimental data. In this respect, Storey et al. supported their cluster analysis by using multi-dimensional scaling (MDS) plotting [92], whilst Aaltonen et al. visualised data using the widely popular multivariate technique, principal component analysis (PCA) [93,94].

PCA is a (bi)linear modelling approach, where the dimensions of the *X*-matrix (corresponding to the number *x*-variables) are reduced into a fewer set of mutually orthogonal (and thus uncorrelated) variables, termed principal components (PCs), best describing the systematic variance of the data. Each PC is the product of samples-score vectors and variables-loading vectors. Plotting two score vectors against each other will give the position of samples (i.e., the rows of the *X*-matrix) in that respective PC direction, while a plot of loading vectors describes the relationship between the original variables (i.e., columns of the *X*-matrix) and the PC direction in question. Combined, the score and loading plots will provide information on how the samples behave mutually [95]. The linearity of PCA means that the modelling and subsequent interpretation is relatively straightforward. In the above-mentioned study [93], NIR data from sulfathiazole samples were decomposed into three PCs, explaining 99.4% cumulative variance. The score plot, showing the position of NIR spectra in these three dimensions, confirmed the presence of three sulfathiazole polymorphs in the sample set. In comparison, solid form screening of nitrofurantoin by combined NIR–Raman required two PCs (94.2% explained variance) for total exploration of the spectral information [94] and revealed two anhydrate forms and one hydrate. Jørgensen et al. monitored the dehydration of erythromycin dihydrate by Raman spectroscopy and PCA [96]. The loading plot provided unique insight into the spectral regions of interest for dehydration phenomena at different temperatures. Another study reports the use of PCA for analysing *in-situ* Raman data from a stable form screening experiment of a disclosed compound [9].

As is the case for all data analysis tools, PCA has its limitations. As such, PCA is a linear modelling approach, and presence of non-linearities in the dataset often requires a higher number of PCs for correct modelling [97]. This was illustrated in a study, where the goal was to visualise the diversity of database containing 218 solvents [58]. Total exploration of the database – in the sense that variance related to all initially selected variables should be described in the model – required five PCs. Clearly, this is not satisfactory from a presentational viewpoint, only permitting two to three dimensions. To accommodate this issue, it can be advantageous to supplement the data analysis with non-linear modelling methods. Self-organising maps (SOMs) in particular, have proven useful for visualising multi-dimensional data in one single plot. SOM fitting is a so-called natural learning algorithm, in which the data are projected onto a map of nodes, while still preserving as much of the topology of the original data as possible. Samples sharing similar properties will be placed in either the same node or in neighbouring regions of the map. Spraggon et al. fitted SOM to image data obtained from a HT crystallisation screen [98]. The SOM made it possible to identify groupings of crystals based on their morphology (unsupervised part), and subsequently for assigning new crys-

tals to the SOM (supervised part). When dealing with multivariate spectral data, it is advisable to reduce the number of *x*-variables in the dataset prior to SOM modelling [97] by, for example, PCA or a genetic algorithm. Ultimately, computational time is decreased, resulting in an increase of overall modelling efficiency. Despite the apparent advantages of SOM, applications in solid form screening are still rare. The lack of literature on this topic might be due to the complexity of natural learning algorithms, which often require several decisions on the settings of algorithmic parameters. For computation of SOMs, the researcher will have to specify parameters such as node distance measure, map size, map shape, neighbourhood function, learning rate, and number of iterations/epochs. These can all affect the specific location of samples on the trained map. Moreover, the random initialisation of nodes (weight-vectors), and the algorithmic approach taken in natural learning mean that slightly varying results are sometimes generated between repeated SOM runs on the same dataset. Evaluation of the quality of the map is therefore crucial. Some previously reported procedures include error estimation either during computation of the SOM [97] or ad-hoc, that is, after computation [58].

2.3.4. Supervised methods

Supervised methods are used for the purpose of predicting a certain sample property. In short, a supervised method is constructed by correlating one or more dependent *x*-variables to at least one independent *y*-variable using different mathematical approaches. The *X*-matrix once again comprises measured data, for example, spectroscopic data, while the *y*-variable designates the sample property that is to be predicted in unknown samples. Supervised methods can roughly be separated into two categories: (1) quantification and (2) classification. Quantification concerns the prediction of a quantitative parameter for unknown samples, usually solid-state purity, while classification aims at assigning unknown samples into user-defined classes. Thus, classification is a qualitative approach. In the current text, focus will be placed on the quantification issue, due to the increased activity within this field resulting in large volumes of information. The reader is referred to excellent textbook material [91] for more information on classification as well as studies by Aldridge et al. [99] and Kogermann et al. [100], in which classification algorithms are applied for API solid form detection. In 2001, Stephenson et al. [101] provided an extensive review on solid form quantification. The literature available can be divided into two categories; (1) in-line quality control during processing and (2) off-line determination of solid-state purity in raw and processed materials or in the final product. In general, there is plenty of information available in the literature on the quantification of crystalline and amorphous forms.

Activity and interest have especially increased within the field of multivariate calibration applied to spectroscopic data (IR, NIR, Raman). In this respect, principal component regression (PCR) and partial least squares (PLS) regression represent useful modelling approaches that utilise either the entire or parts of the spectral region. PCR is in essence a multiple linear regression (MLR) performed on the PCA score values of the samples, the latter representing the systematic variance of the *X*-matrix [91]. Hence, the score values are used for prediction of the desired property (i.e., the *y*-variable) instead of the raw data (intensity values in the case of spectroscopic data). In contrast, the PLS algorithm originates from the fact that the score values are not necessarily predictive for *y*. Thus, the PLS algorithm finds the direction of maximum variance in *X* relevant to the target property, *Y*, thereby creating a new set of scores predictive for *Y* [91]. As a result, a PLS model usually requires a fewer number of components/factors than the corresponding PCR model, and thus may provide a model with better predictive ability. In practice, however, the difference in performance between the two is often small and negligible [91,102]. Indomethacin is a model compound that has undergone much research, concerning the quantification of its solid forms retained in either physical mixtures or in the final dosage form. Otsuka and coworkers have demonstrated how MLR and PCR can be used for multivariate modelling of NIR data of binary mixtures of indomethacin solid forms [103,104]. Same authors used NIR and PCR to predict the amount of γ-indomethacin in tablets containing two different excipients [105]. The calibrated models utilised either the whole or parts of the NIR spectral range. However, using all data points does not necessarily provide optimal models. Instead it is recommended to identify the regions containing the most information with regard to the property being predicted (the *y*-variable). This can greatly improve the accuracy, as well as robustness, of the calibration models. Other studies demonstrate the use of PCR on Raman data for quantification of binary mixtures of polymorphs of carbamazepine [106] and ranitidine HCl [107]. PLS has been widely applied to spectroscopic data for the quantitative determination of API solid forms [102,108–111]. Patel and coworkers compared the predictive abilities of MLR and PLS algorithms to a univariate approach [112], the latter utilising inverse least squares (ILS) regression on a normalised peak. The calibration was performed on second derivative NIR data of binary mixtures containing sulfathiazole Forms I and III for the purpose of quantifying the polymorphic ratio between the two. For this particular dataset, the univariate method provided the lowest predictive error compared to its multivariate counterparts. Same conclusions were drawn in a comparable study concerning the quantification of sulfamethoxazole Form I relative to Form II [113]. Evidently, there are situations in which a simple univariate approach may work better than multivariate analysis. Other studies demonstrate the use of univariate analysis for quantification of polymorphic purity in mixtures by FT-IR [114], and for *in-situ* monitoring of solvent-mediated phase transformation in a batch crystalliser [115] and during dissolution testing [79] by Raman spectroscopy. On an end note, it is relevant to underscore the importance of model performance testing during calibration. For multivariate models, this is often achieved using cross validation (internal) and/or by prediction of an independent test set (external). More information on this topic can be found in reference [91]. Overall, performance testing is a crucial part of multivariate model development where the aim is to obtain reliable predictions of sample properties.

As for unsupervised analysis, non-linear relationship between sample properties and instrumental response may occur. In these cases non-linear natural learning algorithms may produce more accurate models, and thus provide a better alternative to the linear methods. Artificial neural networks (ANNs) for non-linear regression have proven very useful, and may be regarded as the supervised counterpart to the previously discussed self-organising maps (SOMs). Conceptually, ANNs are based on the biological nervous system. From an algorithmic viewpoint, incoming signals (the input) are passed to an artificial neuron body, also denoted as the linear learning machine, where they are weighted and summed. Then they are transformed via a transfer function (or output function) into the outgoing signal (the output). The setting of the weights during the training phase is essential for the proper function of the network, since they determine the relationship between input (*x*-variables, e.g., spectral data) and eventual output (*y*-variable, target property) [91]. The training phase is therefore a crucial part of the ANNs development process. ANNs modelling has previously been used for determining solid-state purity of indomethacin by XRPD [116] and ranitidine hydrochloride by XRPD exclusively [117] or in conjunction with FTIR [118]. The successful application on XRPD data is claimed to be due to the non-linear nature of diffraction intensities as a result of preferred orientation effects. In spite of this, information on the use of ANNs in solid form

screening is still sparse. The complexity of the algorithm and the many user-definable settings are perhaps contributing factors to this fact. Moreover, in spite of offering good predictive properties, interpretation of ANNs models is quite difficult [119].

As is evident from the discussion above, the task of analysing complex solid-state data can be carried out using several unsupervised and supervised techniques. As the techniques are often complementary, it is generally advisable to test more than just a single method. In this context, it must be emphasised that there is a wealth of multivariate techniques that can be applied and whose impact on solid form quantification remains to be elucidated [91]. Undoubtedly, much more research on this topic is to come.

3. Computational methods for prediction of polymorphs

The accurate prediction of a crystal structure and its potential polymorphs is a difficult task [120,121]. There are two reasons for this: firstly, the energies that hold a crystal together are comparatively weak in comparison to intramolecular forces. For example, naphthalene is a solid at room temperature but may be sublimed when heated; this phase change, the sublimation energy, costs 72.4 kJ mol^{-1} of energy [122,123]. Naphthalene contains 10 sp^2-hybridised carbon atoms bonded to each other, these carbon-carbon aromatic bonds have bond energies of the order of 500 kJ mol^{-1} each [124]. The upshot of this very different energy landscape is that computational modelling can provide an accurate picture of nuclear coordinates and *electronic structure* for the molecule naphthalene, but the crystal structure, the intermolecular, is more challenging. With polymorph prediction this difficulty is compounded because the differing polymorphs one may wish to discover *in silico* will have very similar energies. Secondly, the forces that act on molecules to form crystals, and thus determine the lattice energy, are diverse and possess subtle directional properties [121,125,126]. The energetics of molecules in crystals are characterised by a number of terms: (1) Coulombic interactions; (2) Polarisation; (3) Dispersion; (4) Repulsion [120]. To establish the lattice energy each of these terms must be evaluated. One way to model the forces is by considering the atoms of the system and evaluate the atom-atom interactions using a Buckingham-type potential [121] such as

$$E(R_{ij}) = A\exp(-BR_{ij}) - CR_{ij}^{-6} + q_i q_j R_{ij}^{-1} \quad (1)$$

in which the energy between atoms i and j is related to the distance between the atoms (R_{ij}), the charge on each of the atoms (q_i and q_j, respectively) and the parameters A, B and C. The first two terms relate to the dispersion and repulsion forces and the third term to the Coulombic interaction. Attempts to improve the modelling of the lattice energy rely upon getting a quantitatively better picture of the charge distribution about the molecules and the forces that act between molecules. For neutral molecules in a crystal, the dominant energy terms in the lattice energy are the dispersion and repulsion terms which are often 10 times greater than the coulombic term [127].

The prediction of different polymorphs then involves the accurate relative energy determination of the order of 1–2 kJ mol^{-1} [8]. This may seem like an impossible task and is certainly very difficult, but there are a number of successes in the field that have overcome the difficulties alluded to above. The prediction of crystal structures is considered so important that the Cambridge Crystallographic Data Centre (CCDC) has established some blind tests [128–130]. The results of the first two have been discussed by Price [121].

Polymorph Predictor (Molecular Simulations Inc.) is a commercial software that allows one to analyse organic compounds, such as pharmaceuticals, for differing types of polymorphs [131]. This package works by using a Monte Carlo simulated annealing approach, a random search procedure that provides many packing alternates for the compound of interest. The most likely alternates (those with lowest energy) may be optimised by lattice energy minimisation, and these structures ranked in energy terms. The calculated structures may be validated by prediction of the XRPD pattern and this may be compared to experimental data [131]. This method allows each molecule within the unit cell to be adjusted in the minimisation process. This may be done using force fields that include energy values associated with bond stretching, bending and torsional displacement [132]. These force fields are constructed from analysis of structural data for a particular set of compound structures – the parameterisation set. This method can be quite successful if the crystal structure of interest has similar bonding characteristics to the compound structures from which the force field was derived [132]. It is, however, more difficult if the compounds of interest differ significantly from the parameterisation set. For example, there are very good force fields for hydrocarbon interactions [132]. These have been used very effectively in modelling hydrocarbon interactions and structures – but their validity as good force fields in examining aromatic or highly functionalised compounds – which are often found in pharmaceuticals, is less clear [126].

In recent years, two other computational methods have contributed to polymorph prediction and crystal structure modelling of relevance to pharmaceutical systems. The first of these is the DMAREL approach developed by Price and co-workers [121,133,134]. The strategy adopted [72] is to calculate the structure of the molecule of interest using an *ab initio* method, with commercially available software [135]. The charge density of the molecule is then represented by a set of atomic point multipoles, that is, each nuclei is considered to exert a charge that is not simply dependent on R^{-1} (as shown in Eq. (1)), but also includes higher terms, up to R^{-5} [136]. This additional consideration gives a superior Coulombic force field, and thus, in principle, a more accurate evaluation of energies. The dispersion and repulsion terms are provided by parameterised sets [137–139]. Having obtained an accurate model of the charge distribution of the individual molecule, a series of densely packed structures may be obtained, the densest structures are then optimised with respect to the lattice energy using the DMAREL algorithm [136]. It is also possible to predict the phonon modes of crystals and compare these to experimental data, which provides a further test of the validity of the minimised structures [134,140]. The differing conformations that a molecule may have in its crystal form over a calculated gas phase structure have been considered and although not implicitly dealt with in this method there are search strategies to deal with the repacking of flexible molecules [141].

A number of recent studies highlight the success of the DMAREL approach. Progesterone is used as an oral contraceptive. It is available in an optically pure form, termed *nat*-progesterone, which has two polymorphs [142]. The crystal packing landscape of *nat*-progesterone in a number of molecular conformations was explored using the DMAREL strategy [13,143]. As part of this analysis, the mirror image of *nat*-progesterone, *ent*-progesterone, was examined and it was discovered that a racemic mixture of these could pack with a lower energy than either enantiomer. This was experimentally verified by crystallisation of the racemate [13,143].

In a detailed polymorph screening study of carbamazepine [72], it was established that the dimer structural motif that is observed in the 4 known polymorphs is not the only energetically viable motif. An elongated chain structure was also possible. Further to this, the chain structure was obtained experimentally by co-crystallising CBZ with its dihydro analogue (DHC) [144]. The crystal structure of DHC had the requisite chain motif and the co-crystal contained molecules of both DHC and CBZ with this structure.

The PIXEL method, developed by Gavezzotti, is another method for crystal structure prediction [120,125–127,145–148]. In this method, the molecular charge density of the valence electrons is calculated by an *ab initio* method across a three-dimensional volume – with discrete points within the volume being electron-charge pixels. The molecule is then represented by nuclear charges and the electron charge as defined by the pixels within the volume evaluated. The subsequent lattice energies are then determined using the nuclei and electron-charge pixels within the volume of interest. This is a rather appealing strategy because if the charge densities are correctly calculated then the intermolecular forces may be determined. Intermolecular Coulomb energies may be calculated from the nuclei–nuclei, electron pixel–pixel and nuclei electron pixel interactions. The other energies, polarisation and dispersion and repulsion may also be evaluated, although these require the input of some empirical values. This method is not yet accessible to researchers [126] but it has been successful in correlating lattice energies to sublimation energies for almost 100 organic crystals [127]. In addition to simple polymorphs, the issues of differing solvates (or hydrates) and charged species add further complexity to the prediction of structure. The methods described above can certainly deal with this additional complexity, in principle, but the problem becomes more difficult because of the addition of more atoms and, more critically, the fact that the solvate molecules are often smaller than the API and have much shallower interaction energies. This has recently been discussed by Price [149].

In summary, the calculation of crystal structures is challenging, however, progress can be made with careful consideration of the system of interest and judicious choice of force field. Furthermore, a number of new strategies that utilise *ab initio* calculations, and thus should be more generally applicable, are in development.

4. Solid form screening – an industry perspective

In the last decade, big steps have been taken towards the understanding and control of solid forms of APIs. The widespread interest stems not only from the scientific considerations, but also due to the recently emerged regulatory and IP aspects in this field. Both innovator and generic companies have been trying hard to take intellectual gains from the discovery of new solid forms. As a result, solid form patent litigations have become a bottleneck for both sides, which takes many efforts and puts financial burden on the companies [3]. This section provides an overview on the regulatory expectations and IP considerations in solid form screening for both innovator and generic companies.

4.1. Innovator companies

Considering the potential impact of solid forms on product performance, regulatory authorities have put due emphasis on solid form screening and subsequent monitoring. The regulatory guidances have addressed polymorphism in both new drug applications (NDAs) and abbreviated new drug applications (ANDAs) [150]. The ICH Q6A covers solid forms arising during the development of new APIs and decision trees have been provided [6]. Part I of decision tree #4 covers the solid form screening of APIs and their characterisation. However, the extent to which this screen has to be carried out is not answered in this guidance and it is up to the discretion of the pharmaceutical company. Logically, this decision should be based on the solubility, dose and formulation characteristics of the API [151]. Other emphasis given in this guidance is that the proper control and monitoring should be placed whenever multiple solid forms are present, and specification with respect to polymorphic purity should be incorporated if necessary. The specification on polymorphic purity is very important if the API has solubility-limited bioavailability and/or is prone to solid-state transformations during processing and/or storage. Therefore, once a new API is chosen for potential development, it is imperative to screen for the solid forms it may possess, and to identify the most suitable one with respect to solubility and stability as early as possible. However, there is no method that can provide absolute confidence that the ideal solid form has been obtained, and sometimes the final form used in the product may indeed arise at the later stages of development [14,152]. Atorvastatin was initially formulated as an amorphous salt during the development phase. However, it is reported that during Phase III clinical studies, the salt crystallised and its properties changed, which compelled the developer, Warner–Lambert, to conduct additional bridging studies to demonstrate acceptability of the new product relative to that used in the pivotal registration studies [14]. These kinds of events are costly, consuming more development time as well as resources. As such innovator companies are free to choose any suitable solid form as long as there are no IP issues involved, which is highly unlikely during early stage as knowledge regarding the API would normally have remained within the company itself. The knowledge generated by conducting solid form screening can provide the innovator company an opportunity to build a patent portfolio around different forms and therefore a means to enhance product lifecycle management [153]. A convergence of events negatively affecting pharmaceutical product lifecycle has elevated the value of solid form IP, thus making comprehensive solid form screening early in the development very important. Innovator companies have tried to protect solid forms by patents and have gained extra years for the product beyond the expiry of the basic molecule patent. One of the earliest cases of this kind was that of ranitidine hydrochloride, in which GSK had patent protection for form II even though the basic molecule patent had expired [154,155]. Although the generic companies were able to launch products with form I, the form II patent helped postpone the generic entry. Table 6 shows a few more examples where branded pharmaceutical products from the innovator companies have patent protection due to solid form(s) beyond the expiry of the basic molecule patent. These data are taken from the US Food and Drug Administration's (FDA) Orange Book database and there can be other solid form patents for these molecules which are not present in the Orange Book [156]. Innovators are required by FDA regulations to identify patents claiming API, formulation, or specific therapeutic use of the new pharmaceutical in connection with the drug approval process. The patents covering solid forms are also part of this requirement. By including patent in the Orange Book, the innovator gets the advantage of 30 months stay if ANDA is filed with paragraph IV certification for any of the listed patents, and thus, can delay the generic entry. However, the NDA applicant or holder is required to submit a patent claiming a different polymorph from that described in the NDA if a product containing the new polymorph will perform the same as the product described in the NDA with respect to dissolution, solubility, and bioavailability [157]. This means that for any solid form other than the one used in the NDA, the innovator has to prove that the new solid form has the same therapeutic effect by conducting bioequivalence studies. But even if the innovator does not include patent(s) related to other solid forms in the Orange Book and does not gain from the 30 months stay, they can still get advantages of patent protection.

In a hypothetical case where the solid form patent is the only limiting factor for the generic entry, generic firms can launch their product if they can discover new a solid form which does not have IP protection and has suitable characteristics for product development. This was the case when Teva found a way around Merck's patents on its crystalline form of alendronate (the active ingredient in the blockbuster Fosamax®) and was able to launch generic version much earlier [158]. Thus, by patenting a maximum number of

Table 6
List of some molecules having solid form patent(s) listed in the Orange Book [156]

Generic name	Brand name	Basic molecule patent (expiry date)	Solid form patent (expiry date)	Solid form patented
Clopidogrel bisulphate	Plavix®	US 4847265 (Nov 17, 2011)	US 6429210 (Jun 10, 2019) US 6504030 (Jun 10, 2019)	Polymorph II
Atorvastatin	Lipitor®	US 5273995 (Dec 28, 2010)	US 5969156 (Jul 08, 2016)	Polymorph I, II, IV and their hydrates
Olanzapine	Zyprexa®	US 5605897 (Feb 25, 2014)	US 5736541 (Mar 24, 2015) US 6251895 (Sep 23, 2017)	Polymorph II, Dihydrate D
Fexofenadine	Allegra®	US 5578610 (Nov 26, 2013)	US 7135571 (May 18, 2014) US 7138524 (May 18, 2014)	Anhydrate Form I, Hydrate Form II
Donepezil Hydrochloride	Aricept®	US 4895841 (Nov 25, 2010)	US 5985864 (Dec 30, 2016) US 6140321 (Dec 30, 2016) US 6245911 (Dec 01, 2018)	Polymorph II, III, IV, V and A, B, C
Gatifloxacin	Zymar®	US 4980470 (Dec 15, 2009)	US 5880283 (Dec 05, 2015)	Sesquihydrate
Ziprasidone hydrochloride	Geodon®	US 4831031 (Mar 02, 2012)	US 5312925 (Sep 01, 2012)	Monohydrate
Gabapentin	Neurontin®	US 4024175 (May 17, 1994)	US 4894476 (May 02, 2008)	Monohydrate

possible solid forms, innovators can block this route, even though these solid forms would not be used in the pharmaceutical product. The case where an alternative solid form patented by another company has better characteristics can also have major implications. An example of such is topiramate sodium where J&J ended up licensing and paying royalties for the trihydrate form developed and patented by Transform Pharmaceuticals [159]. After the launch of a dosage form by the innovator, a new solid form can help development of novel drug delivery methods with different release profiles and routes of administration. This is particularly important for APIs whose poor solubility is the main hurdle in the development. The above discussion exemplifies the importance of solid form screening for innovator companies. Extensive solid form screening can not only provide them with scientific advantages, but also help meet regulatory requirements and maximise returns from drug development by means of IP. It can be inferred from Table 6 that solid form patents can provide extra protection anywhere between 1 and 9 years after the basic molecule patent has expired. In case of IP gain, the timing of filing the form patents is important which should be after the core new chemical entity (NCE) patent filing, but before a competitor has the opportunity to perform solid form screening. The timing of solid form patent filing also depends on the properties of the molecule and other kinds of patent protection such as the formulation and the method of use present in the patent portfolio [160]. Hence, for innovator companies it is logical to carry out solid form screening either before clinical trials or early in clinical trials to maximise the benefit to drug development and reduce IP risk.

4.2. Generic companies

W.C. McCrone's statement regarding polymorphism, "*the number of forms known for a given compound is proportional to the time and energy spent in research on that compound*", becomes quite evident when looking at the contribution of generic pharmaceutical companies in the discovery of new solid forms of APIs. Once the branded pharmaceutical product reaches the market and its API has shown business potential, generic companies start putting effort into solid form screening. Compared to the innovator companies, the generic companies are able to invest more in chemical development and process chemistry simply because they can rely on much of the other work already done by the innovator. As innovators are protecting more and more solid forms by patents, solid form screening by generic companies becomes imperative for an early as possible launch of a generic product. For a single branded pharmaceutical product there are many generic players working around. Hence, time and energy spent in solid form screening becomes manifold, fulfilling the stated requirement for the discovery of new forms. The aspect of polymorphism in generic product development is well covered in recent publications and regulatory guidances [150,161,162]. FDA issued its latest guidance on polymorphism-related issues to be considered for ANDA submission in 2007 [150]. This guidance is more elaborate compared to ICH Q6A and provides decision trees for the development of a generic product. Decision tree #1 provides recommendations on when specifications for polymorphic form(s) for the API and/or the pharmaceutical product may be appropriate. Polymorphs are unlikely to have a significant effect on bioavailability when all the forms have the same apparent solubilities or all the forms are highly soluble. The guidance puts emphasis on solid form screening to get knowledge about all the solid forms that the API may have along with the use of published literature and patents. Decision tree #2 gives an approach for setting specifications for polymorphs when at least one form is known to have low solubility based on the BCS. Decision tree #3 provides an approach when considering whether to set specifications for polymorphs in the pharmaceutical product. Generally, specifications for polymorphs in pharmaceutical products are not necessary if the most thermodynamically stable polymorphic form is used or if the same form is used in an approved product of the same dosage form. The recommendations regarding which solid forms to be considered for monitoring and control are given based on the physicochemical properties of the API. Moreover, it has clarified the issue of "sameness" in ANDAs, in which the guidance mentions that the different polymorphic forms do not render the substances different APIs for the purpose of ANDA approvals. Here, the term 'polymorphic forms' includes polymorphs, amorphous form, solvates and hydrates. Over the years, FDA has approved many ANDAs having the API in a different solid form from the one in the respective Reference Listed Drug (RLD). Therefore, if a solid form present in a RLD is still under patent protection and the basic molecule patent has expired, generic companies can develop their products with another solid form, provided it meets other requirements of FDA for ANDA approval. Generic companies are working hard to tap such opportunities, and competition has forced them to carry out solid form screening as early as possible. In the US, the first ANDA approved with paragraph IV certification (i.e., particular Orange Book patent(s) is invalid, unenforceable, or will not be infringed) is entitled to 180 days marketing exclusivity [163]. In many cases where the solid form patent is listed in the Orange Book, generic companies have opted to file ANDA with a new solid form developed by their own solid form screening which provides an opportunity for paragraph IV certification. So by carrying out thorough screening, generic companies can not only develop their own generic products with new solid forms, but they can also block other generic launches with the help of marketing exclusivity and patent protection. One notable example here is sertraline hydrochloride, which is the active ingredient in the blockbuster antidepressant Zoloft®. Teva filed ANDA with paragraph IV certification on Orange Book listed patent US 5248699 which claims sertraline polymorph, and

eventually gained 180 days marketing exclusivity. Teva secured patent protection on its novel solid forms by several patents and later went on to sue many other generic companies who had filed ANDAs for sertraline hydrochloride. In fact, this race of being the first generic has created a situation where there are more solid forms patented by generic companies compared to innovators. Overall, it is prudent for the generic companies to carry out extensive solid form screening to get the best possible advantages from new solid forms.

5. Conclusions

The development of a new pharmaceutical requires a deep understanding of solid-state phenomena. In order for a pharmaceutical product to succeed, the possible existence and performance of various solid forms need to be thoroughly investigated and the IP strategies carefully considered before commercial launch. Therefore, solid form screening has become an essential part of pharmaceutical development and product lifecycle management.

The number of new pharmaceuticals is decreasing, and the molecules that reach the later phases of drug development are becoming ever more structurally complex, which creates further challenges for solid form screening. New biomacromolecular compounds are often administered as solutions. However, the solid state is, and will remain, the most stable state of matter and for this reason, a very attractive option for formulation. In the future, research should focus on not only optimisation of the crystalline material, but also exploring the means for stabilisation of amorphous formulations for increased solubility, and thus bioavailability.

The tools available for solid form screening have evolved radically during the past decade – we can now explore the solid forms computationally, perform thousands of experimental crystallisations with miniaturised high-throughput screening technologies, and identify new solid phases fast. We should, however, pay special attention on the evaluation part of the screening. Understanding of basic thermodynamics together with robust design of experiments and powerful data analysis are the keys to successful solid form screening.

Acknowledgement

The Academy of Finland, Finnish Cultural Foundation and Finnish Pharmaceutical Society are acknowledged for funding (J.A.). Professor Thomas Rades is thanked for fruitful discussions.

References

[1] J. Haleblian, W. McCrone, Pharmaceutical applications of polymorphism, J. Pharm. Sci. 58 (1969) 911–929.
[2] J. Thun, L. Seyfarth, J. Senker, R.E. Dinnebier, J. Breu, Polymorphism in Benzamide: solving a 175-year-old riddle, Angew. Chem., Int. Ed. 46 (2007) 6729–6731.
[3] J. Bernstein, Polymorphism and patents, Polymorphism in Molecular Crystals, Clarendon Press, Oxford, 2002. pp. 297–307.
[4] J. Bulow, The gaming of pharmaceutical patents, in: A.B. Jaffe, J. Lerner, S. Stern (Eds.), Innovation Policy and the Economy 4, The MIT Press, Cambridge, MA, 2004, pp. 145–185.
[5] S.R. Chemburkar, J. Bauer, K. Deming, H. Spiwek, K. Patel, J. Morris, R. Henry, S. Spanton, W. Dziki, W. Porter, J. Quick, P. Bauer, J. Donaubauer, B.A. Narayanan, M. Soldani, D. Riley, K. McFarland, Dealing with the impact of Ritonavir polymorphs on the late stages of bulk drug process development, Org. Process Res. Dev. 4 (2000) 413–417.
[6] ICH Harmonised Tripartite Guideline Q6A, Test procedures and acceptance criteria for new drug substances and new drug products: chemical substances, International Conference on Harmonisation, 1999.
[7] R.J. Davey, Crystallization: how come you look so good?, Nature 428 (2004) 374–375
[8] J.D. Dunitz, Are crystal structures predictable?, Chem Commun. (2003) 545–548.
[9] J.M. Miller, B.M. Collman, L.R. Greene, D.J.W. Grant, A.C. Blackburn, Identifying the stable polymorph early in the drug discovery-development process, Pharm. Dev. Tech. 10 (2005) 291–297.
[10] C.-H. Gu, V. Young, D.W.J. Grant, Polymorph screening: influence of solvents on the rate of solvent-mediated polymorphic transformation, J. Pharm. Sci. 90 (2001) 1878–1890.
[11] D. Singhal, W. Curatolo, Drug polymorphism and dosage form design: a practical perspective, Adv. Drug Del. Rev. 56 (2004) 335–347.
[12] W.I. Cross, N. Blagden, R.J. Davey, R.G. Pritchard, M.A. Neumann, R.J. Roberts, R.C. Rowe, A whole output strategy for polymorph screening: combining crystal structure prediction, graph set analysis, and targeted crystallization experiments in the case of diflunisal, Cryst. Growth Des. 3 (2003) 151–158.
[13] R.W. Lancaster, P.G. Karamertzanis, A.T. Hulme, D.A. Tocher, D.F. Covey, S.L. Price, Racemic progesterone: predicted in silico and produced in the solid state, Chem. Commun. (2006) 4921–4923.
[14] C.R. Gardner, C.T. Walsh, Ö. Almarsson, Drugs as materials: valuing physical form in drug discovery, Nat. Rev. Drug Discov. 3 (2004) 926–934.
[15] S.R. Byrn, R.R. Pfeiffer, J.G. Stowell, Polymorphs, Solid-State Chemistry of Drugs, SSCI Inc., West Lafayette, 1999. pp. 143–231.
[16] D.W.J. Grant, Theory and origin of polymorphism, in: H.G. Brittain (Ed.), Polymorphism in Pharmaceutical Solids, Marcel Dekker Inc., New York, 1999, pp. 1–33.
[17] E.V. Boldyreva, T.P. Shakhtshneider, H. Ahsbahs, H. Sowa, H. Uchtmann, Effect of high pressure on the polymorphs of paracetamol, J. Therm. Anal. Calorim. 68 (2002) 437–452.
[18] R.J. Davey, N. Blagden, G.D. Potts, R. Docherty, Polymorphism in molecular crystals: stabilization of a metastable form by conformational mimicry, J. Am. Chem. Soc. 119 (1997) 1767–1772.
[19] P. Vishweshwar, J.A. McMahon, J.A. Bis, M.J. Zaworotko, Pharmaceutical co-crystals, J. Pharm. Sci. 95 (2006) 499–516.
[20] S.R. Byrn, R.R. Pfeiffer, J.G. Stowell, Hydrates and solvates, Solid-State Chemistry of Drugs, SSCI Inc., West Lafayette, 1999.
[21] A.L. Gillon, N. Feeder, R.J. Davey, R. Storey, Hydration in molecular crystals – a Cambridge Structural Database analysis, Cryst. Growth Des. 3 (2003) 663–673.
[22] K.R. Morris, Structural aspects of hydrates and solvates, in: H.G. Brittain (Ed.), Polymorphism in Pharmaceutical Solids, Marcel Dekker Inc., New York, 1999, pp. 125–181.
[23] B.C. Hancock, G. Zografi, Characteristics and significance of the amorphous state in pharmaceutical systems, J. Pharm. Sci. 86 (1997) 1–12.
[24] N. Blagden, R.J. Davey, R. Rowe, R. Roberts, Disappearing polymorphs and the role of reaction by-products: the case of sulphathiazole, Int. J. Pharm. 172 (1998) 169–177.
[25] L. Yu, Amorphous pharmaceutical solids: preparation, characterization and stabilization, Adv. Drug Del. Rev. 48 (2001) 27–42.
[26] A.M. Kaushal, P. Gupta, A.K. Bansal, Amorphous drug delivery systems: molecular aspects, design, and performance, Crit. Rev. Ther. Drug Carrier Syst. 21 (2004) 133–193.
[27] B.C. Hancock, M. Parks, What is the true solubility advantage for amorphous pharmaceuticals?, Pharm Res. 17 (2000) 397–404.
[28] T.L. Threlfall, Analysis of organic polymorphs. A review, Analyst 120 (1995) 2435–2460.
[29] G.L. Amidon, H. Lennernäs, V.P. Shah, J.R. Crison, A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability, Pharm. Res. 12 (1995) 413–420.
[30] P. Kahela, R. Aaltonen, E. Lewing, M. Anttila, E. Kristoffersson, Pharmacokinetics and dissolution of two crystalline forms of carbamazepine, Int. J. Pharm. 14 (1983) 103–112.
[31] M.C. Meyer, A.B. Straughn, E.J. Jarvi, G.C. Wood, F.R. Pelsor, V.P. Shah, The bioinequivalence of carbamazepine tablets with a history of clinical failures, Pharm. Res. 9 (1992) 1612–1616.
[32] P. Di Martino, A.M. Guyot-Hermann, P. Conflant, M. Drache, J.C. Guyot, A new pure paracetamol for direct compression: the orthorhombic form, Int. J. Pharm. 128 (1996) 1–8.
[33] G. Nichols, C.S. Frampton, Physicochemical characterization of the orthorhombic polymorph of paracetamol crystallized from solution, J. Pharm. Sci. 87 (1998) 684–693.
[34] E. Suihko, V.-P. Lehto, J. Ketolainen, E. Laine, P. Paronen, Dynamic solid-state and tabletting properties of four theophylline forms, Int. J. Pharm. 217 (2001) 225–236.
[35] C. Sun, D.J.W. Grant, Influence of crystal structure on the tableting properties of sulfamerazine polymorphs, Pharm. Res. 18 (2001) 274–280.
[36] S.R. Byrn, W. Xu, A.W. Newman, Chemical reactivity in solid-state pharmaceuticals: formulation implications, Adv. Drug Del. Rev. 48 (2001) 115–136.
[37] S.R. Byrn, P.A. Sutton, B. Tobias, J. Frye, P. Main, Crystal structure, solid-state NMR spectra, and oxygen reactivity of five crystal forms of prednisolone *tert*-butylacetate, J. Am. Chem. Soc. 110 (1988) 1609–1614.
[38] Y. Guo, S.R. Byrn, G. Zografi, Physical characteristics and chemical degradation of amorphous quinapril hydrochloride, J. Pharm.Sci. 89 (2000) 128–143.
[39] R.J. Davey, J. Garside, From Molecules to Crystallizers. An Introduction to Crystallization, Oxford Science Publications, Oxford, 1999.
[40] J.W. Mullin, Industrial techniques and equipment, in: Crystallization, Butterworth Heinemann, Oxford, 2001, pp. 315–402.
[41] M. Volmer, Die chemische reaktion. Bd. IV. Kinetik der phasenbildung, Steinkopf, Leipzig, 1939.

[42] S. Khoskhoo, J. Anwar, Crystallization of polymorphs: the effect of solvent, J. Phys. D: Appl. Phys. 26 (1993) 90–93.
[43] J. Bernstein, R.J. Davey, J.-O. Henck, Concominant polymorphs, Angew. Chem., Int. Ed. 38 (1999) 3440–3461.
[44] R. Hilfiker, S.M. De Paul, M. Szelagiewcz, Approaches to polymorphism screening, in: R. Hilfiker (Ed.), Polymorphism in the Pharmaceutical Industry, Wiley-VCH Verlag GmbH, Weinheim, 2006, pp. 287–308.
[45] J.K. Guillory, Generation of polymorphs, hydrates, solvates and amorphous solids, in: H.G. Brittain (Ed.), Polymorphism in Pharmaceutical Solids, Marcel Dekker Inc., New York, 1999, pp. 183–226.
[46] S.L. Morissette, Ö. Almarsson, M.L. Peterson, J.F. Remenar, M.J. Read, A.V. Lemmo, S. Ellis, M.J. Cima, C.R. Gardner, High-throughput crystallization: polymorphs, salts, co-crystals and solvates of pharmaceutical solids, Adv. Drug Del. Rev. 56 (2004) 275–300.
[47] S. Mirza, I. Miroshnyk, J. Heinämäki, L. Christiansen, M. Karjalainen, J. Yliruusi, Influence of solvents on the variety of crystalline forms of erythromycin, AAPS PharmSci 5 (2003) (Article 12).
[48] I. Miroshnyk, L. Khriachtchev, S. Mirza, J. Rantanen, J. Heinämäki, J. Yliruusi, Insight into thermally induced phase transformations of erythromycin A dihydrate, Cryst. Growth Des. 6 (2006) 369–374.
[49] S.L. Childs, L.J. Chyall, J.T. Dunlap, D.A. Coates, B.C. Stahly, G.P. Stahly, A metastable polymorph of metformin hydrochloride: isolation and characterization using capillary crystallization and thermal microscopy techniques, Cryst. Growth Des. (2004) 441–449.
[50] A.S. Myerson, B.A. Garetz, Method for using laser light to control crystal form, U.S. Patent 6426406, 2002.
[51] G. Ruecroft, D. Hipkiss, T. Ly, N. Maxted, P.W. Cains, Sonocrystallization: the use of ultrasound for improved industrial crystallization, Org. Process Res. Dev. 9 (2005) 923–932.
[52] N. Rodríguez-Hornedo, S.J. Nehm, A. Jayasankar, Cocrystals: design properties and formation mechanisms, in: J. Swarbrick (Ed.), Encyclopedia of Pharmaceutical Technology, third ed., Informa Healthcare USA, Inc., New York, 2007, pp. 615–635.
[53] G.P. Stahly, Diversity in single- and multiple-component crystals. The search for and prevalence of polymorphs and cocrystals, Cryst. Growth Des. 7 (2007) 1007–1026.
[54] B.Y. Shekunov, P. York, Crystallization processes in pharmaceutical technology and drug delivery design, J. Cryst. Growth 211 (2000) 122–136.
[55] M. Kitamura, K. Nakamura, Effects of solvent composition and temperature on polymorphism and crystallization behavior of thiazole-derivative, J. Cryst. Growth 236 (2002) 676–686.
[56] M. Kitamura, Controlling factor of polymorphism in crystallization process, J. Cryst. Growth 237–239 (2002) 2205–2214.
[57] C.-H. Gu, H. Li, R.B. Gandhi, K. Raghavan, Grouping solvents by statistical analysis of solvent property parameters: implication to polymorph screening, Int. J. Pharm. 283 (2004) 117–125.
[58] M. Allesø, F. van den Berg, C. Cornett, F.S. Jørgensen, B. Halling-Sørensen, H. Lopez de Diego, L. Hovgaard, J. Aaltonen, J. Rantanen, Solvent diversity in polymorph screening, J. Pharm. Sci. 97 (2008) 2145–2159.
[59] M. Allesø, J. Rantanen, J. Aaltonen, C. Cornett, F. van den Berg, Solvent subset selection for polymorph screening, J. Chemom., in press, doi:10.1002/cem.1107.
[60] D. Xu, N. Redman-Furey, Statistical cluster analysis of pharmaceutical solvents, Int. J. Pharm. 339 (2007) 175–188.
[61] L.-F. Huang, W.-Q. Tong, Impact of solid state properties on developability assessment of drug candidates, Adv. Drug Del. Rev. 56 (2004) 321–334.
[62] W. Beckmann, Seeding the desired polymorph: background, possibilities, limitations, and case studies, Org. Process Res. Dev. 4 (2000) 372–383.
[63] C.-H. Gu, K. Chatterjee, V. Young, D.J.W. Grant, Stabilization of a metastable polymorph of sulfamerazine by structurally related additives, J. Cryst. Growth 235 (2002) 471–481.
[64] X. He, J.G. Stowell, K.R. Morris, R.R. Pfeiffer, H. Li, G.P. Stahly, S.R. Byrn, Stabilization of a metastable polymorph of 4-methyl-2-nitroacetanilide by isomorphic additives, Cryst. Growth Des. 1 (2001) 305–312.
[65] A.L. Grzesiak, M. Lang, K. Kim, A.J. Matzger, Comparison of the four anhydrous polymorphs of carbamazepine and the crystal structure of form I, J. Pharm. Sci. 92 (2003) 2260–2271.
[66] M. Lang, J.W. Kampf, A.J. Matzger, Form IV of carbamazepine, J. Pharm. Sci. 91 (2002) 1186–1190.
[67] A. Park, L.J. Chyall, J. Dunlap, C. Schertz, D. Jonaitis, B.C. Stahly, S. Bates, R. Shipplett, S. Childs, New solid-state chemistry technologies to bring better drugs to market: knowledge-based decision making, Expert Opin. Drug Discov. 2 (2007) 145–154.
[68] P.J. Desrosiers, The potential of preform, Mod. Drug Discov. (2004) 40–43.
[69] S.L. Morissette, S. Soukasene, D. Levinson, M.J. Cima, Ö. Almarsson, Elucidation of crystal form diversity of the HIV protease inhibitor Ritonavir by high-throughput crystallization, Proc. Natl. Acad. Sci. USA 100 (2003) 2180–2184.
[70] M.L. Peterson, S.L. Morissette, C. McNulty, A. Goldsweig, P. Shaw, M. LeQuesne, J. Monagle, N. Encina, J. Marchionna, A. Johnson, J. Gonzalez-Zugasti, A.V. Lemmo, S.J. Ellis, M.J. Cima, Ö. Almarsson, Iterative high-throughput polymorphism studies on acetaminophen and an experimentally derived structure for form III, J. Am. Chem. Soc. 124 (2002) 10958–10959.
[71] Ö. Almarsson, M.B. Hickey, M.L. Peterson, S.L. Morissette, S. Soukasene, C. McNulty, M. Tawa, J.M. MacPhee, J.F. Remenar, High-throughput surveys of crystal form diversity of highly polymorphic pharmaceutical compounds, Cryst. Growth Des. 3 (2003) 927–933.
[72] A.J. Florence, A. Johnston, S.L. Price, H. Nowell, A.R. Kennedy, N. Shankland, An automated parallel crystallization search for predicted crystal structures and packing motifs of carbamazepine, J. Pharm. Sci. 95 (2006) 1918–1930.
[73] H.G. Brittain, Physical Characterization of Pharmaceutical Solids, Marcel Dekker Inc., New York, 1995.
[74] R. Hilfiker, Polymorphism: In the Pharmaceutical Industry, Wiley-VCH Verlag GmbH KGaA, Weinheim, 2006.
[75] H.G. Brittain, Polymorphism in Pharmaceutical Solids, Marcel Dekker Inc., New York, 1999.
[76] D. Giron, M. Mutz, S. Garnier, Solid-state of pharmaceutical compounds, J. Thermal Anal. Calorim. 77 (2004) 709–747.
[77] R.A. Lane, G. Buckton, The novel combination of dynamic vapour sorption gravimetric analysis and near infra-red spectroscopy as a hyphenated technique, Int. J. Pharm. 207 (2000) 49–56.
[78] A.D. Gift, L.S. Taylor, Hyphenation of Raman spectroscopy with gravimetric analysis to interrogate water–solid interactions in pharmaceutical systems, J. Pharm. Biomed. Anal. 43 (2007) 14–23.
[79] J. Aaltonen, P. Heinänen, L. Peltonen, H. Kortejärvi, V.P. Tanninen, L. Christiansen, J. Hirvonen, J. Yliruusi, J. Rantanen, In situ measurement of solvent-mediated phase transformations during dissolution testing, J. Pharm. Sci. 95 (2006) 2730–2737.
[80] A.M. Dyas, U.U. Shah, Dissolution and dissolution testing, in: J. Swarbrick (Ed.), Encyclopedia of Pharmaceutical Technology, third ed., vol. 2, Informa Healthcare USA, Inc., New York, 2007.
[81] D.C. Montgomery, Design and Analysis of Experiments, Wiley, New York, 2001.
[82] L.X. Yu, R.A. Lionberger, A.S. Raw, R. D'Costa, Applications of process analytical technology to crystallization process, Adv. Drug Del. Rev. 56 (2004) 349–369.
[83] N. Al-Zoubi, K. Kachrimanis, S. Malamataris, Effects of harvesting and cooling on crystallization and transformation of orthorhombic paracetamol in ethanolic solution, Eur. J. Pharm. Sci. 17 (2002) 13–21.
[84] M. Sheikhzadeh, S. Murad, S. Rohani, Response surface analysis of solution-mediated polymorphic transformation of buspirone hydrochloride, J. Pharm. Biomed. Anal. 45 (2007) 227–236.
[85] M. Fujiwara, Z.K. Nagy, J.W. Chew, R.D. Braatz, First-principles and direct design approaches for the control of pharmaceutical crystallization, J. Process Contr. 15 (2005) 493–504.
[86] R. Bro, M. Jakobsen, Exploring complex interactions in designed data using GEMANOVA. Color changes in fresh beef during storage, J. Chemom. 16 (2002) 294–304.
[87] S. Lowry, D. Dalrymple, A. Song, V. Rosso, C. Pommier, J. Venit, Integrating a Raman microscope into the workflow of a high-throughput crystallization laboratory, J. Ass. Lab. Automat. 11 (2006) 75–84.
[88] J.F. Remenar, J.M. MacPhee, B.K. Larson, V.A. Tyagi, J.H. Ho, D.A. McIlroy, M.B. Hickey, P.B. Shaw, O. Almarsson, Salt selection and simultaneous polymorphism assessment via high-throughput crystallization: the case of sertraline, Org. Process Res. Dev. 7 (2003) 990–996.
[89] R. Hilfiker, J. Berghausen, F. Blatter, A. Burkhard, S. De Paul, B. Freiermuth, A. Geoffroy, U. Hofmeier, C. Marcolli, B. Siebenhaar, M. Szelagiewicz, A. Vit, M. von Raumer, Polymorphism – integrated approach from high-throughput screening to crystallization optimization, J. Thermal Anal. Calorim. 73 (2003) 429–440.
[90] S. Datta, D.W.J. Grant, Crystal structures of drugs: advances in determination, prediction and engineering, Nat. Rev. Drug Discov. 3 (2004) 42–57.
[91] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics – Part B, Elsevier, Amsterdam, 1998.
[92] R. Storey, R. Docherty, P. Higginson, C. Dallman, C. Gilmore, G. Barr, W. Dong, Automation of solid form screening procedures in the pharmaceutical industry – how to avoid the bottlenecks, Crystallogr. Rev. 10 (2004) 45–56.
[93] J. Aaltonen, J. Rantanen, S. Siiria, M. Karjalainen, A. Jorgensen, N. Laitinen, M. Savolainen, P. Seitavuopio, M. Louhi-Kultanen, J. Yliruusi, Polymorph screening using near-infrared spectroscopy, Anal. Chem. 75 (2003) 5267–5273.
[94] J. Aaltonen, C.J. Strachan, K. Pollanen, J. Yliruusi, J. Rantanen, Hyphenated spectroscopy as a polymorph screening tool, J. Pharm. Biomed. Anal. 44 (2007) 477–483.
[95] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemom. Intell. Lab. Syst. 2 (1987) 37–52.
[96] A.C. Jørgensen, I. Miroshnyk, M. Karjalainen, K. Jouppila, S. Siiria, O. Antikainen, J. Rantanen, Multivariate data analysis as a fast tool in evaluation of solid state phenomena, J. Pharm. Sci. 95 (2006) 906–916.
[97] B.K. Lavine, C.E. Davidson, D.J. Westover, Spectral pattern recognition using self-organizing maps, J. Chem. Inf. Comp. Sci. 44 (2004) 1056–1064.
[98] G. Spraggon, S.A. Lesley, A. Kreusch, J.P. Priestle, Computational analysis of crystallization trials, Acta Crystallogr., Sect. D 58 (2002) 1915–1923.
[99] P.K. Aldridge, C.L. Evans, H.W.I. Ward, S.T. Colgan, N. Boyer, P.J. Gemperline, Near-IR detection of polymorphism and process-related substances, Anal. Chem. 68 (1996) 997–1002.
[100] K. Kogermann, J. Aaltonen, C.J. Strachan, K. Pollänen, P. Veski, J. Heinämäki, J. Yliruusi, J. Rantanen, Qualitative in situ analysis of multiple solid-state forms using spectroscopy and partial least squares discriminant modeling, J. Pharm. Sci. 96 (2007) 1802–1820.

[101] G.A. Stephenson, R.A. Forbes, S.M. Reutzel-Edens, Characterization of the solid state: quantitative issues, Adv. Drug Del. Rev. 48 (2001) 67–90.
[102] H.Y. Aboul-Enein, A.A. Bunaciu, S. Fleschin, Analysis of mebendazole polymorphs by Fourier transform IR spectrometry using chemometric methods, Biopolymers 67 (2002) 56–60.
[103] M. Otsuka, F. Kato, Y. Matsuda, Comparative evaluation of the degree of indomethacin crystallinity by chemoinfometrical Fourier-transformed near-infrared spectroscopy and conventional powder X-ray diffractometry, AAPS PharmSci 2 (2000) (Article 9).
[104] M. Otsuka, F. Kato, Y. Matsuda, Determination of indomethacin polymorphic contents by chemometric near-infrared spectroscopy and conventional powder X-ray diffractometry, Analyst 126 (2001) 1578–1582.
[105] M. Otsuka, F. Kato, Y. Matsuda, Y. Ozaki, Comparative determination of polymorphs of indomethacin in powders and tablets by chemometrical near-infrared spectroscopy and X-ray powder diffractometry, AAPS PharmSciTech 4 (2003) (Article 19).
[106] C.J. Strachan, D. Pratiwi, K.C. Gordon, T. Rades, Quantitative analysis of polymorphic mixtures of carbamazepine by Raman spectroscopy and principal components analysis, J. Raman Spectrosc. 35 (2004) 347–352.
[107] D. Pratiwi, J.P. Fawcett, K.C. Gordon, T. Rades, Quantitative analysis of polymorphic mixtures of ranitidine hydrochloride by Raman spectroscopy and principal components analysis, Eur. J. Pharm. Biopharm. 54 (2002) 337–341.
[108] J.W. Braga, R.J. Poppi, Figures of merit for the determination of the polymorphic purity of carbamazepine by infrared spectroscopy and multivariate calibration, J. Pharm. Sci. 93 (2004) 2124–2134.
[109] K.R. Lee, G. Zuber, L. Katrincic, Chemometrics approach to the determination of polymorphism of a drug compound by infrared spectroscopy, Drug Dev. Ind. Pharm. 26 (2000) 135–147.
[110] K. Pöllänen, A. Häkkinen, M. Huhtanen, S.P. Reinikainen, M. Karjalainen, J. Rantanen, M. Louhi-Kultanen, L. Nyström, DRIFT-IR for quantitative characterization of polymorphic composition of sulfathiazole, Anal. Chim. Acta 544 (2005) 108–117.
[111] A. Heinz, M. Savolainen, T. Rades, C. Strachan, Quantifying ternary mixtures of different solid-state forms of indomethacin by Raman and near-infrared spectroscopy, Eur. J. Pharm. Sci. 32 (2007) 182–192.
[112] A.D. Patel, P.E. Luner, M.S. Kemper, Low-level determination of polymorph composition in physical mixtures by near-infrared reflectance spectroscopy, J. Pharm. Sci. 90 (2001) 360–370.
[113] A.D. Patel, P.E. Luner, M.S. Kemper, Quantitative analysis of polymorphs in binary and multi-component powder mixtures by near-infrared reflectance spectroscopy, Int. J. Pharm. 206 (2000) 63–74.
[114] R. Helmy, G.X. Zhou, Y.W. Chen, L. Crocker, T. Wang, R.M. Wenslow Jr., A. Vailaya, Characterization and quantitation of aprepitant drug substance polymorphs by attenuated total reflectance fourier transform infrared spectroscopy, Anal. Chem. 75 (2003) 605–611.
[115] J. Schöll, D. Bonalumi, L. Vicum, M. Mazzotti, M. Muller, In situ monitoring and modeling of the solvent-mediated polymorphic transformation of L-glutamic acid, Cryst. Growth Des. 6 (2006) 881–891.
[116] T. Okumura, M. Nakazono, M. Otsuka, K. Takayama, An accurate quantitative analysis of polymorphs based on artificial neural networks, Colloids Surf. B 49 (2006) 153–157.
[117] S. Agatonovic-Kustrin, V. Wu, T. Rades, D. Saville, I.G. Tucker, Powder diffractometric assay of two polymorphic forms of ranitidine hydrochloride, Int. J. Pharm. 184 (1999) 107–114.
[118] S. Agatonovic-Kustrin, T. Rades, V. Wu, D. Saville, I.G. Tucker, Determination of polymorphic forms of ranitidine-HCl by DRIFTS and XRPD, J. Pharm. Biomed. Anal. 25 (2001) 741–750.
[119] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics – Part A, Elsevier, Amsterdam, 1998.
[120] A. Gavezzotti, Ten years of experience in polymorph prediction: what next?, CrystEngComm 4 (2002) 343–347
[121] S.L. Price, The computational prediction of pharmaceutical crystal structures and polymorphism, Adv. Drug Del. Rev. 56 (2004) 301–319.
[122] D. Williams, Coulombic interactions in crystalline hydrocarbons, Acta Crstallogr. Sect. A 30 (1974) 71–77.
[123] R.S. Bradley, T.G. Cleasby, The vapour pressure and lattice energy of some aromatic ring compounds, J. Chem. Soc. (1953) 1690–1692.
[124] P.W. Atkins, Physical Chemistry, Oxford University Press, Oxford, 1992.
[125] A. Gavezzotti, A solid-state chemist's view of the crystal polymorphism of organic compounds, J. Pharm. Sci. 96 (2007) 2232–2241.
[126] J.D. Dunitz, A. Gavezzotti, Molecular recognition in organic crystals: directed intermolecular bonds or nonlocalized bonding?, Angew Chem., Int. Ed. 44 (2005) 1766–1787.
[127] A. Gavezzotti, Quantitative ranking of crystal packing modes by systematic calculations on potential energies and vibrational amplitudes of molecular dimers, J. Chem. Theory Comput. 1 (2005) 834–840.
[128] 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 Crystallogr. Sect. B 58 (2002) 647–661.
[129] 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, A test of crystal structure prediction of small organic molecules, Acta Crystallogr. Sect. B 56 (2000) 697–714.
[130] 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 Crystallogr. Sect. B 61 (2005) 511–527.
[131] P.H. Young, H.Y. Ando, Analysis of known crystals to design polymorph prediction strategies, J. Pharm. Sci. 96 (2007) 1203–1236.
[132] S.L. Mayo, B.D. Olafson, W.A.I. Goddard, DREIDING: a generic force field for molecular simulations, J. Phys. Chem. 94 (1990) 8897–8909.
[133] T.C. Lewis, D.A. Tocher, G.M. Day, S.L. Price, A computational and experimental search for polymorphs of parabanic acid – a salutary tale leading to the crystal structure of oxo-ureido-acetic acid methyl ester, CrystEngComm 5 (2003) 3–9.
[134] G.M. Day, S.L. Price, M. Leslie, Atomistic calculations of phonon frequencies and thermodynamic quantities for crystals of rigid organic molecules, J. Phys. Chem. B 107 (2003) 10919–10933.
[135] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, Gaussian, Inc., Pittsburgh, PA, 2003.
[136] D.J. Willock, S.L. Price, M. Leslie, C.R.A. Catlow, The relaxation of molecular-crystal structures using a distributed multipole electrostatic model, J. Comput. Chem. 16 (1995) 628–647.
[137] D.E. Williams, Improved intermolecular force field for molecules containing H, C, N, and O atoms, with application to nucleoside and peptide crystals, J. Comput. Chem. 22 (2001) 1154–1166.
[138] D.E. Williams, Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon, J. Mol. Struct. 486 (1999) 321–347.
[139] D.E. Williams, Improved intermolecular force field for crystalline oxohydrocarbons including OH–O hydrogen bonding, J. Comput. Chem. 22 (2001) 1–20.
[140] G.M. Day, J.A. Zeitler, W. Jones, T. Rades, P.F. Taday, Understanding the influence of polymorphism on phonon spectra: lattice dynamics calculations and terahertz spectroscopy of carbamazepine, J. Phys. Chem. B 110 (2006) 447–456.
[141] P.G. Karamertzanis, S.L. Price, Energy minimization of crystal structures containing flexible molecules, J. Chem. Theory Comput. 2 (2006) 1184–1199.
[142] H. Campsteyn, L. Dupont, O. Dideberg, Structure cristalline et moléculaire de la progestérone, $C_{21}H_{30}O_2$, Acta Crystallogr. B 28 (1972) 3032–3042.
[143] R.W. Lancaster, P.G. Karamertzanis, A.T. Hulme, D.A. Tocher, T.C. Lewis, S.L. Price, The polymorphism of progesterone: stabilization of a 'disappearing' polymorph by co-crystallization, J. Pharm. Sci. 96 (2007) 3419–3431.
[144] A.J. Florence, C.K. Leech, N. Shankland, K. Shankland, A. Johnston, Control and prediction of packing motifs: a rare occurrence of carbamazepine in a catemeric configuration, CrystEngComm 8 (2006) 746–747.
[145] A. Gavezzotti, Calculation of intermolecular interaction energies by direct numerical integration over electron densities. I. Electrostatic and polarization energies in molecular crystals, J. Phys. Chem. B 106 (2002) 4145–4154.
[146] C.J. Eckhardt, A. Gavezzotti, Computer simulations and analysis of structural and energetic features of some crystalline energetic materials, J. Phys. Chem. B 111 (2007) 3430–3437.
[147] A. Gavezzotti, Hierarchies of intermolecular potentials and forces: progress towards a quantitative evaluation, Struct. Chem. 16 (2005) 177–185.
[148] A. Gavezzotti, Modeling hydrogen bonded crystals, J. Mol. Struct. 615 (2002) 5–12.
[149] S.L. Price, From crystal structure prediction to polymorph prediction: interpreting the crystal energy landscape, Phys. Chem. Chem. Phys. 10 (2008) 1996–2009.
[150] Guidance for industry. ANDAs: pharmaceutical solid polymorphism. Chemistry, manufacturing and controls information, U.S. Food and Drug Administration, 2007.
[151] S.P.F. Miller, A.S. Raw, L.X. Yu, Scientific considerations of pharmaceutical solid polymorphism in regulatory applications, in: R. Hilfiker (Ed.), Polymorphism, Wiley-VCH Verlag GmBH & Co. KGaA, Weinheim, Germany, 2006, pp. 385–403.
[152] A. Llinas, J.M. Goodman, Polymorph control: past, present and future, Drug Discov. Today 13 (2008) 198–210.
[153] J. Lucas, P. Burgess, When form equals substance: the value of form screening in product life-cycle management, Pharma Voice (2004) 54–57.
[154] B.J. Price, J.W. Clitherow, J. Bradshaw, Aminoalkyl furan derivatives, U.S. Patent 4128658, 1978.

[155] D.L. Crookes. Aminoalkyl furan derivative, U.S. patent 4521431, 1985.
[156] Electronic Orange Book: Approved Drug Products with Therapeutic Equivalence Evaluations, U.S. Food and Drug Administration, <http://www.fda.gov/cder/ob/default.htm>, accessed on 1st February, 2008.
[157] Applications for FDA approval to market a new drug: patent submission and listing requirements and application of 30-month stays on approval of abbreviated new drug applications certifying that a patent claiming a drug is invalid or will not be infringed; Final rule, Fed. Regist. 68 (2003) 36675–36712.
[158] H. Slowik, The battle for IP, In Vivo 21 (2003) 75–84.
[159] M. Stuart, Transform, J&J, topiramate: new life, old products, In Vivo 22 (2004) 71–72.
[160] L. Howard, Use of patents in drug lifecycle management, J. Generic Med. 4 (2007) 230–236.
[161] D.A. Snider, W. Addicks, W. Owens, Polymorphism in generic drug product development, Adv. Drug Del. Rev. 56 (2004) 391–395.
[162] A.S. Raw, M.S. Furness, D.S. Gill, R.C. Adams, F.O. Holcombe Jr., L.X. Yu, Regulatory considerations of pharmaceutical solid polymorphism in Abbreviated New Drug Applications (ANDAs), Adv. Drug Del. Rev. 56 (2004) 397–414.
[163] 180-Day generic drug exclusivity under the Hatch–Waxman amendments to the federal food, drug, and cosmetic act, U.S. Food and Drug Administration, 1998.