Pharmaceutical Digital Design: From Chemical Structure through Crystal Polymorph to Conceptual Crystallization Process

Christopher L. Burcham[†], Michael F. Doherty[††], Baron G. Peters[‡], Sarah L. Price[*], Matteo Salvalaglio[**], Susan M. Reutzel-Edens[†a], Louise S. Price[*], Ravi Kumar Reddy Addula[‡], Nicholas Francia[**b], and Vikram Khanna[††]

† Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN 46285

†† Department of Chemical Engineering, University of California, Santa Barbara, Santa Barbara, CA 93106

‡ Department of Chemical and Biomolecular Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801

* Department of Chemistry, University College London, 20 Gordon Street, London, UK WC1H 0AJ

** Department of Chemical Engineering, University College London, London, UK WC1E 6BT

[a] Present address: SuRE Pharma Consulting, LLC, Zionsville, IN 46077

[b] Present address: Cambridge Crystallographic Data Centre, Cambridge, UK CB2 1EZ

17 Synopsis: This work considers the theoretical basis of crystal structure prediction (CSP), free

18 energy, solubility, morphology and growth rate prediction, and the current state of nucleation

19 simulation to provide the conceptual process design for industrial crystallizations of

20 pharmaceutical compounds. This is illustrated by applying the modeling techniques to real

21 examples, olanzapine and succinic acid. We describe and demonstrate the promise of using *ab*

22 *initio* computer modeling for solid form selection and process design in pharmaceutical

23 development from only a molecular structure.

Table of Contents

3

69   ABSTRACT: A workflow for the digital design of crystallization processes starting from the

70   chemical structure of the active pharmaceutical ingredient (API) is a multi-step, multi-disciplinary

71   process. A simple version would be to first predict the API crystal structure and from it the

72   corresponding properties of solubility, morphology, and growth rates, assume that the nucleation

73   would be controlled by seeding, and then use these parameters to design the crystallization process.

74   This is usually an over-simplification as most APIs are polymorphic, and the most stable crystal

75   of the API alone may not have the required properties for development into a drug product. This

76   perspective, from the experience of a Lilly Digital Design project, considers the fundamental

77   theoretical basis of crystal structure prediction (CSP), free energy, solubility, morphology and

78   growth rate prediction, and the current state of nucleation simulation. This is illustrated by

79   applying the modeling techniques to real examples, olanzapine and succinic acid. We demonstrate

80   the promise of using *ab initio* computer modeling for solid form selection and process design in

81   pharmaceutical development. We also identify open problems in the application of current

82   computational modeling and achieving the accuracy required for immediate implementation that

83   are currently limiting the applicability of the approach.

1   Introduction

The development of a new pharmaceutical product begins with a hypothesis that a new molecule will either promote or interrupt a biochemical pathway to affect a disease state. The target molecule is tested in clinical trials for safety (Phase I) and efficacy (Phase II and III) and ultimately progresses to the commercialization of a new medicine if successful. The journey for a molecule to become a medicine is long, usually fraught with many obstacles, and very expensive. The average cost to develop a new drug has been estimated to be between $2.3 billion (reported in 2023[1, 2]) and $3 billion (reported in 2013[3]), with the probability that a Phase I compound will successfully progress to product approval at just under 12%.[3]

Innovative pharmaceutical companies have long been utilizing computational tools to improve the success rate of identifying prospective drug candidates to validate *in vivo*. Computational tools are regularly used in the initial phase of drug discovery, e.g. to predict medicinal chemistry targets and pathways,[4-6] to identify candidate molecules, to compute binding affinity,[7, 8] to forecast the metabolism of a drug[9] and to anticipate toxicological issues.[10, 11] These computational efforts increasingly drive the subsequent experimental efforts by identifying promising candidates prior to synthesis and *in vivo* testing.

Simulation tools hold similar potential to shorten the time and lower the cost of product design and development, but *ab initio* simulations have not been used to the same extent in the commercial phase of product and process development. Except for perhaps the use of pharmacokinetic and pharmacodynamic (PK-PD) models to predict drug absorption for solid oral dosage forms[12-15] and retrosynthetic techniques for route selection,[16-19] development has remained mostly empirically and experimentally driven, with computational approaches taking a supporting role as a drug enters the product formulation and commercial process design stage.

107   The paths envisaged to digitally design a drug product, i.e., to progress a drug through clinical

108   trials and develop it into a commercial drug product, are depicted in Figure 1. Overall, the process

109   is not linear and there are many opportunities for iteration. However, many key milestones cannot

110   be achieved until a precursor activity is completed – and for solid oral dosage forms, nothing is

111   more critical than identifying the crystal form of the drug to carry into commercial development.

112   Should the selection of the crystal form change, much of the product development process must

113   be repeated, including clinical trials to demonstrate the equivalence of the new form to that

114   previously tested in humans.

115
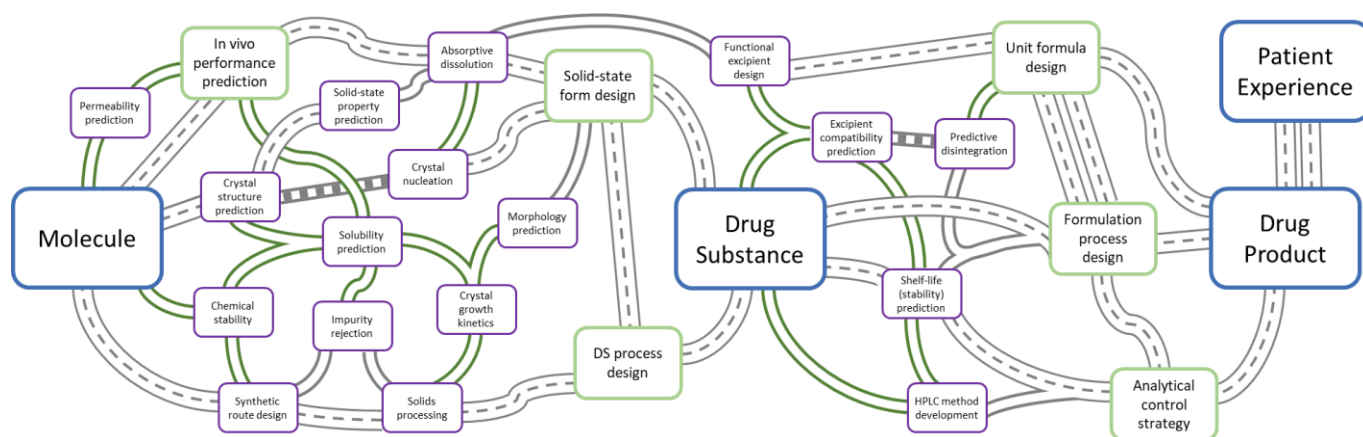
116

117   Figure 1.   Roadmap illustrating the major (in bold) elements and associated work streams

118   necessary for developing a new molecular entity into a solid oral drug product.   (DS is drug

119   substance).

120   The concept of digital design and utilization of computer modeling to complement the

121   experimental input required for the design of solid forms and the crystallization processes to make

122   them has emerged over the past two decades,[20-22] and is the subject of a recent review.[23] There has

123   been much work assessing progress in each step, from the series of blind tests of crystal structure

124   prediction organized by the Cambridge Crystallographic Data Centre (CCDC)[24] and the aqueous

125   solubility prediction challenges,[25] to Faraday Discussion meetings on crystallization,[26] as well the

126   output of the crystallization working group of the Enabling Technologies Consortium.[22, 27]

127   Considerable progress is also being made with the use of informatics, for example, the extensive

128   range of tools in the CCDC's suite of programs, to complement experimental and computational

129   chemistry efforts.[28]

130   This contribution evaluates how current computational tools can be combined in a workflow to

131   design crystallization processes from the molecular structure of the API. It is not intended as a

132   review of the different steps, but rather a **demonstration** of the use of *ab initio* computational tools

133   in the development cycle of a pharmaceutical product through the **integration** of the current state

134   of the art in selecting the desired solid state form of the API, through calculation of the requisite

135   physical properties necessary for optimal bioavailability and downstream processability. This

136   paper seeks to outline the fundamental physical basis of predicting each property and illustrate

137   these calculations on two systems, olanzapine, an atypical antipsychotic agent originally marketed

138   by Lilly as Zyprexa©,[29] and succinic acid. Both are relatively small molecules compared with

139   current small molecule APIs under development. Consideration is given to both pure computation

140   of absolute or relative properties and where the input of some experimental data into the simulation

141   can provide a wide range of data of the required accuracy.

142   Inevitably, the practical computational models are more suited to certain types of molecules than

143   others, and the range of molecules to which they have been applied successfully with the accuracy

144   necessary for digital design varies. This leads to an outline of the open problems, both for each

145   step and overall, towards developing digital design strategies that could be deployed for the

146   crystallization of a wide range of pharmaceuticals, including multicomponent systems.

147   2   The Vision of *Ab Initio* Crystallization Process Design

148   Most drugs are developed as solid dosage forms, such as tablets or capsules,[30] due to patient

149 preference for orally administered drugs. *In silico* modeling should, in principle, provide a forward

150 look at the effort needed to develop an API into a new solid oral dosage form. Such methods

151 should not only help to design a product that will reliably deliver the drug to its target, but also

152 help to estimate and manage complexity, direct the experimental effort in an efficient and material

153 sparing way, optimize resources, and minimize the risk of downstream failures. Herein, we share

154 our vision for how *ab initio* methods will continue to evolve towards enabling crystallization

155 process design.

156   Development generally progresses along two streams, one to produce the drug substance from

157 commercially available starting materials and the other to produce the final dosage form with the

158 desired attributes (size, dose, shape, etc.). The two streams are not mutually exclusive given that

159 the process to make the drug product must be able to accommodate the physical properties (bulk

160 density, powder flowability, etc.) of the drug substance.[31] Moreover, a coordinated effort may be

161 required to optimize the drug substance alongside the formulation to achieve the desired *in vivo*

162 performance of the solid oral dosage form.

163   For either situation, the choice of a suitable crystalline form of the API is one that meets the

164 needs of both drug substance and drug product development. This means that in practice,

165 concurrent with selecting a commercially viable synthetic route and before other key activities can

166 commence, the desired solid form of the drug substance is identified. Thus, the proposed digital

167 first workflow depicted in Figure 2 starts from the molecular structure with the prediction of the

168 static crystal structure landscape (CSP_0) to identify low energy structures that are plausible

169 polymorphs. These structures are then refined through the determination of the free energy at

170 room temperature to allow for prediction of phase diagrams (as a function of temperature, relative

171   humidity (RH), and sometimes pressure) as well as modeling and prediction of key properties

172   (solubility, etc.) to support the selection of the solid form.
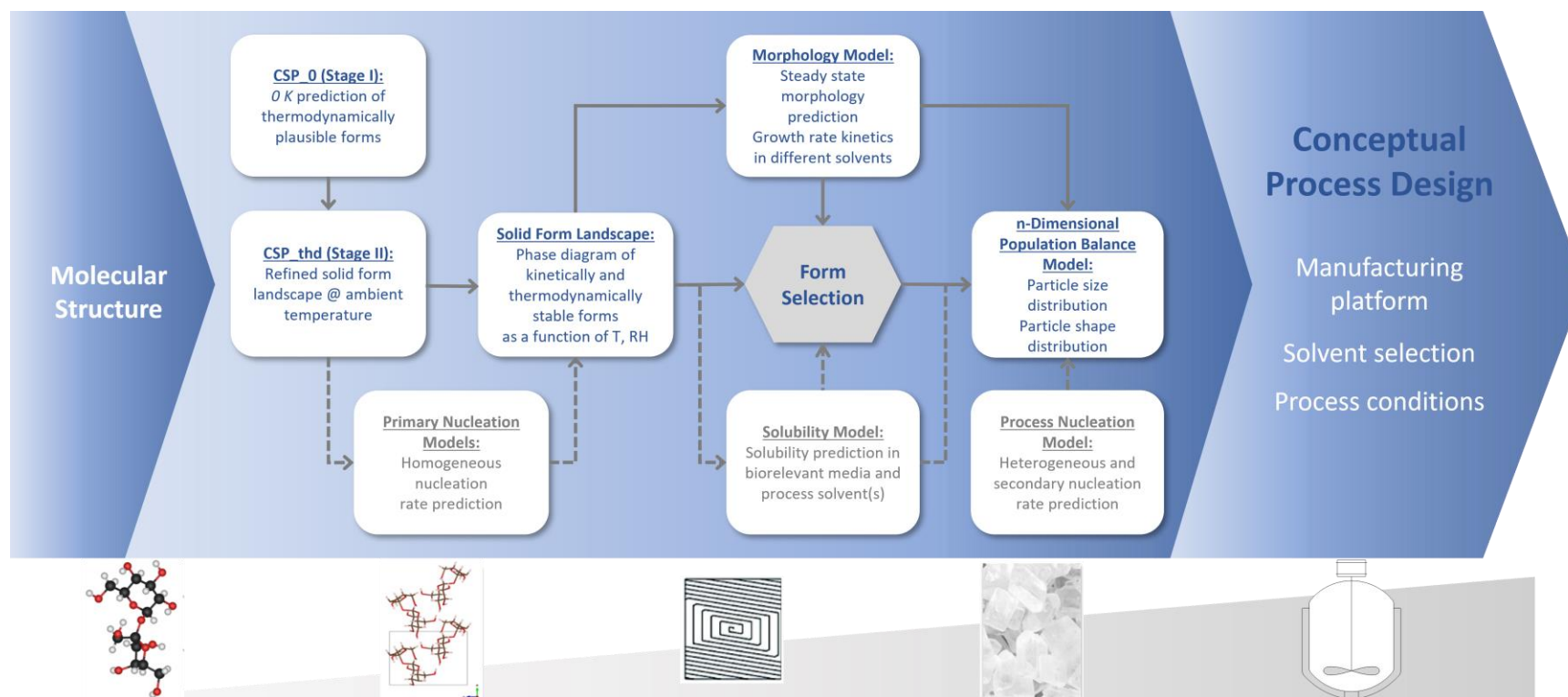
Figure 2. Schematic digital workflow to define a conceptual crystallization process, starting from molecular structure on the left, and proceeding to the crystallizer on the right. *Ab initio* models demonstrated in this paper are depicted in blue; those for which first-principles models are not yet sufficiently accurate (currently determined experimentally) are colored gray.

177   Once the desired crystal form is selected, the design of an API crystallization process typically

178   commences by considering multiple objectives: purification/chemical purity, physical properties,

179   product yield, robustness, and productivity. Some models already exist to design crystallization

180   processes conceptually, helping to identify feasible operating conditions, guide economic

181   assessments, and in conjunction with control strategies, achieve target particle sizes and crystal

182   morphologies prior to the start of an experimental development program. Current models,

183   however, require existing experimental data (solubility, growth rates, nucleation rates) as inputs.

184   For example, the fate and purge of impurities after crystallization can be estimated based on the

185   impurity partition coefficient between the solvent and crystalline drug substance, and solubility of

186   the crystalline drug and impurity,[32] although a computational impurity rejection model based on

187   impurities forming solid solutions has been recently proposed.[33]

188   In our proposed digital workflow, polymorph-specific and facet-specific growth rates, together

189   with their dependence on solvent composition, and product solubility, are computed for putative

190   polymorphs (thermodynamically competitive crystal structures), then used in process models to

191   predict the crystal shape and to provide insight to the process design. We employ state-of-the-art

192   multi-dimensional population balance modeling to predict particle size and shape distributions,

193   but for these to be predicted by growth rates alone, we consider only seeded batch processes,

194   operating under conditions such that primary and secondary nucleation are minimized. While

195   reliable empirical models for secondary nucleation do exist, they require a few material- and

196   crystallizer-specific parameters, which must be experimentally determined. Thus, without a means

197   to predict secondary nucleation kinetics from a first-principles approach (see Section 6), the fully

198   *ab initio* conceptual design of a continuous crystallization process[34] is not yet possible.

199   The main components of the proposed digital workflow in Figure 2 (blue boxes) can, in

200   principle, be applied to any small molecule to design and ultimately deliver a drug substance with

201   desirable physical properties. However, situations may arise where additional models (not shown)

202   are needed to inform the crystallization process design. For example, content uniformity models[35-

203   37] may be required for high potency, low dose drugs and bioavailability models[13, 14] are necessary

204   for BCS class II/IV molecules to establish the upper limit of the crystal size distribution that the

205   process must achieve. Whether the individual models are foundational (broadly applicable to small

206   molecules) or situational, they must be compatible for the overall workflow to succeed. That is,

207   the accuracy of the outputs must be sufficiently high to be useful as inputs to subsequent models.

208   To this end, we draw attention herein to three key areas (gray boxes in Figure 2) where accurate

209   first-principles models are needed for the vision of conceptual (batch or continuous) crystallization

210   process design to be realized: primary nucleation (for polymorph discovery), biorelevant

211   solubility, and secondary nucleation.

212

213   3   Crystal Polymorph Selection

214   Following the expensive and very public recall of Norvir® from the market due to the sudden

215   appearance of a more stable, less soluble polymorph of the API, ritonavir, in the drug product,[38-

216   41] all large pharmaceutical companies have made major investments in solid form and salt

217   screening programs to mitigate the risk of late-appearing forms. The selection of the solid form

218   usually starts with screening studies to identify viable solid compositions from potentially many

219   salts (if ionizable), cocrystals, or the parent compound itself. This is followed by the screening

220   and ultimately selection of the desired crystal polymorph (neat form or hydrate) among the

221   crystalline 'hits'. Theoretically, these exercises are best done holistically, but in reality, the

222   process of selecting a solid form is iterative, limited by the supply of the drug substance and the

223   time allotted for form screening.

224   Getting form selection right is non-negotiable as it affects downstream operations, as well as *in*

225   *vivo* performance. However, while industry has adopted a number of 'best practices' over the years

226   to rapidly screen diverse crystallization conditions, the path to a commercially-viable crystal form

227   remains very unpredictable and molecule-dependent; there is no one-size-fits-all recipe to

228   crystallize a molecule for the first time, let alone in a solid form that is suitable for a commercial

229   drug product. Furthermore, form selection is often conducted at a time in development when the

230   certainty of the compound moving forward into commercialization or even further into Phase II or

231   Phase III clinical trials is extremely low.[42] Therein lies the conundrum; how much investment is

232   required to determine that a suitable composition (parent compound, salt, cocrystal) and the

233   preferred, most stable polymorph have been identified, when the commercial success rate is so low

234   and when speed to the clinic is critical? The need to get the crystal form right, preferably the first

235   time to minimize rework, must be balanced with the cost of screening a potentially large portfolio

236   of molecules, most of which will never progress further in development. This is where the *in silico*

237   design of solid forms, starting with the prediction of the crystal structure, has the potential to

238   change the game for solid oral dosage form development.

239

240   3.1   Crystal Structure Prediction

241   The pharmaceutical industry is interested in crystal structure prediction (CSP) from the desire to

242   right-size the search for the possible crystal forms (i.e. the polymorphs of the neat API[43] and at

243   least its hydrates[44]). The main risk to avoid is the late appearance of a form more stable than the

244   one under development, as this may lead to its "disappearance"[39] or a sudden need to change the

13

245 form, the manufacturing process or storage conditions. As a complement to the experimental

246 screening and characterization of solid forms,[45] a CSP study can provide confidence that the most

247 stable form is known, and, in favorable cases, aid the design of an appropriate experimental search

248 for that form, as well as identify metastable forms to design around during process development.

249 The ideal CSP computational code would predict all the polymorphs that could be

250 experimentally realized and give a recipe for obtaining the first sample of each polymorph. This is

251 indeed the ultimate aim,[46] but the series of blind tests of CSP organized by the CCDC[24] show that

252 this is still an aspiration in this rapidly developing area. Currently, the first stage in the type of

253 CSP that is most commonly applied in industry,[47] referred to as CSP_0, is the search for structures

254 that are the most stable minima in the lattice energy. This is the energy required to separate a

255 (hypothetical) static infinite perfect crystal into infinitely separated molecules in their lowest

256 energy conformation, approximating the relative stability at 0 K. Since the relative stability of

257 polymorphs often changes with temperature and pressure, we need to develop the calculation of a

258 crystal energy landscape at processing and storage conditions (CSP_thd). Other thermodynamic

259 factors, such as the balance of bulk and surface energies, reflecting particle size should be taken

260 into account[48] as this can affect the relative stability of polymorphs, and can lead to the observation

261 of new polymorphs in confined crystallization experiments.[49] Environmental factors, such as

262 water activity (or relative humidity), must also be considered, such that the crystal structures of

263 anhydrates and hydrates of different stoichiometries can be directly compared on the same energy

264 landscape.[50]

265 A major limitation of CSP_0 is that it usually generates significantly more crystal structures

266 within the likely energy range of polymorphism than are found experimentally.[51, 52] Although some

267 structures may appear as disorder components in observed polymorphs, many of these structures

14

268   may be artifacts of the neglect of the temperature-dependent molecular motions within the crystals.

269   Other structures identified as putative polymorphs could be kinetically forbidden, because of the

270   kinetics of nucleation and growth relative to the ability to transform to a more stable form. Despite

271   the advances in modeling nucleation and growth kinetics at the atomic scale,[53] our current

272   understanding of the kinetic competition involved in apparent polymorph stability is not yet

273   sufficiently mature to be encapsulated into a CSP workflow.

274   3.1.1   Lattice energy CSP

275   Current CSP_0 methods search for the possible crystal structures corresponding to a given

276   molecular diagram, i.e. covalent connectivity. Hence, typical CSP_0 searches will not include

277   crystal structures containing different tautomers, or allow a cocrystal to convert to a salt, even

278   when these are more stable. CSP studies consider a fixed stoichiometry of multi-component forms,

279   such as cocrystals, solvates and salts. The relative energies from CSP studies with different

280   tautomers or compositions can be compared, but the current success of CSP_0 comes from the

281   cancellation of errors in relative lattice energies, and the introduction of different types of

282   molecules and additional types of intermolecular interactions means that direct comparison is more

283   prone to error. The searches are generally limited to a specified number of crystallographically

284   independent molecules (often just $Z'=1$) and range of space groups. This is often leveraged to

285   concentrate the computational search into specific, chemically relevant, regions of the crystal

286   packing space. For example, a search can be restricted to the chiral space groups if the API is

287   chiral, and the other enantiomer cannot appear in the crystal structure.

288   A CSP search must also consider the range of conformations that could plausibly be of

289   sufficiently low energy to appear in a crystal structure. Typically many tens of thousands, even

290   millions of structures are generated for a pharmaceutical API with only one molecule in the

291 asymmetric unit cell (Z′=1). Searches with two or more independent molecules in the unit cell are

292 required for multicomponent systems and when there is experimental evidence that the API has

293 crystallized with two or more conformations or packing environments, and these CSP studies

294 generate many more structures than a (Z′=1) search. This number of crystal structures practically

295 necessitates a hierarchical ranking process to find the most stable distinct crystal structures, using

296 increasingly accurate and expensive methods of optimizing the structures to the nearest minimum

297 in the lattice energy, while progressively eliminating structures that are too high in energy to be

298 plausible polymorphs or are duplicates. The cost of a worthwhile CSP_0 strongly depends on the

299 molecule: the diversity of functional groups determines the range of possible intermolecular

300 interactions that dictate the crystallization process, and the conformational flexibility determines

301 how much the molecule can be distorted from the isolated molecule conformation by the packing

302 forces. It also depends on the range of crystal structures covered in the structure generation stage.

303 Hence, it is difficult to estimate in advance how many low energy crystal structures the CSP_0

304 will generate within the energy window being considered for possible polymorphism, and how

305 accurate the relative energies need to be to give a reliable order of stability. This ranges from

306 clearly monomorphic molecules, where there is only one way to pack the molecules densely in all

307 three dimensions, to molecules that have many closely related structures generated by CSP_0 that

308 contribute to the disordered crystal structures often observed in experiments.[51, 54]

309   An efficient ranking of crystal structures necessitates computationally inexpensive methods for

310 the initial calculation of lattice energies. Traditional point-charge classical force fields that have

311 been extensively developed for liquid state and biomolecular simulations are rarely sufficiently

312 accurate for the final stages of CSP_0, as shown by the early blind tests of CSP. Molecule-specific

313 force fields can be generated by purpose-developed applications (e.g. GRACE[55]) for use in the

16

314   structure generation stage, but these differ from traditional force fields by dispensing with the

315   assumption that the same atomic types and parameters can be used for intermolecular and

316   intramolecular forces. An approach that recognizes that the molecule is very similar in its crystal

317   structures uses the definition $E_{latt} = U_{inter} + \Delta E_{intra}$, where U_{inter} is the sum over all the

318   intermolecular interactions within the crystal as calculated with an anisotropic atom-atom force

319   field, and ΔE_{intra} is the penalty for conformational change. This approach is denoted Ψ_{mol}, as it

320   only requires electronic structure calculations on the isolated molecule.[56] The molecular electronic

321   structures calculations provide the molecular structure within the crystal, the conformational

322   energy penalty ΔE_{intra} for changes from the most stable isolated molecule conformation, and the

323   atomic multipole moments required to model the electrostatic contribution to U_{inter}. Many

324   worthwhile CSP_0 studies combine this anisotropic atom-atom electrostatic model with an

325   empirical *exp-6* repulsion dispersion potential,[57, 58] and are very inexpensive for small rigid

326   molecules, where only one molecular wavefunction needs to be calculated. The success of this

327   approach can be attributed to the atomic dipoles and quadrupoles modeling the anisotropic

328   electrostatic effects of the lone pairs and π electron density, hence capturing the directionality of

329   hydrogen-bonding, $\pi\cdots\pi$ stacking and other dominant intermolecular interactions.[59, 60] The theory

330   of intermolecular forces allows the development of increasingly realistic models for the

331   intermolecular interactions from the molecular charge distribution, and the generation of non-

332   empirical models.[61] It has recently been shown that improving the quality of the *ab initio* method

333   used to describe the intramolecular energies (ΔE_{intra}) can significantly increase the accuracy of the

334   Ψ_{mol} approach.[62] However, because the charge distribution of the molecule changes with

335   conformation, the cost of this approach scales with the number of conformations that have to be

336   considered, although various interpolation and database methods mean that this approach can be
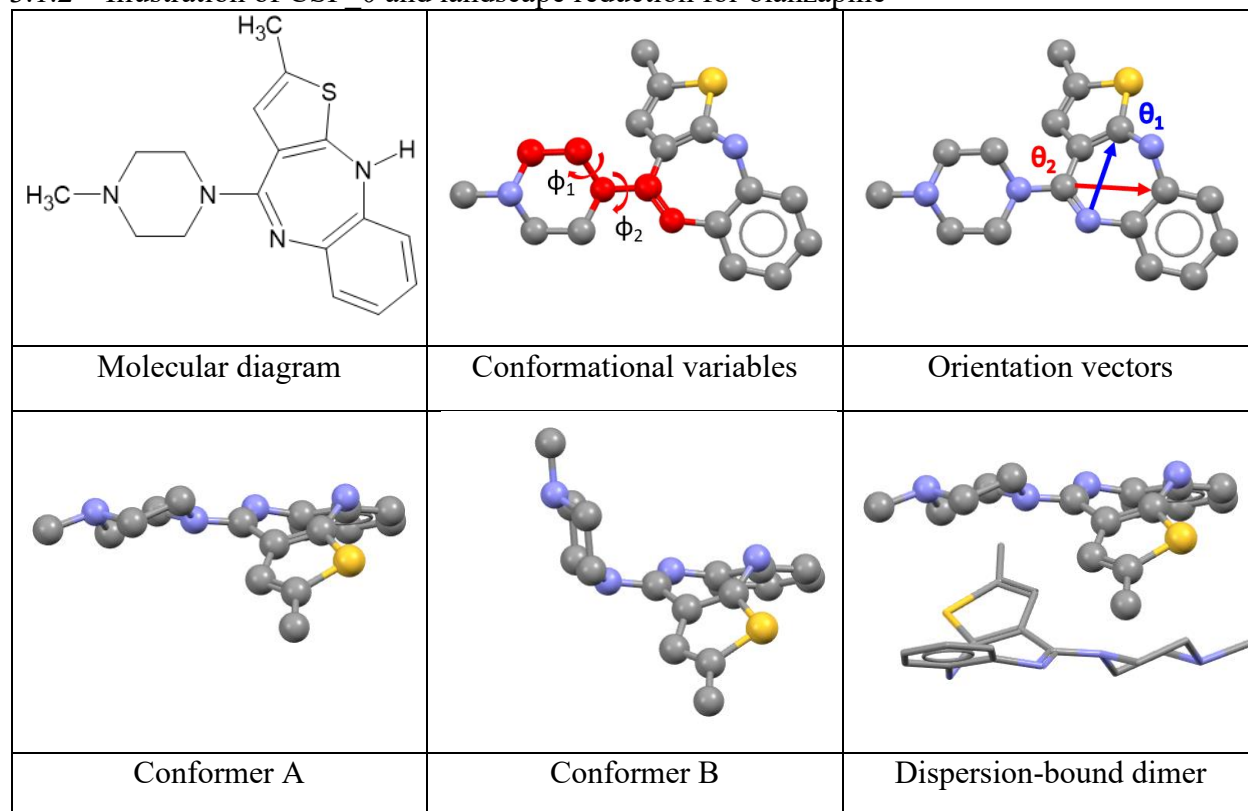
337 applied to pharmaceuticals.[63] A disadvantage of this method is that it is difficult to accurately

338 calculate the additional stabilization of the lattice by changes in the molecular charge distribution

339 (i.e. polarization) induced by the surrounding molecules.[58] This can be approximated in a structure-

340 independent fashion by calculating the molecular charge density (and hence ΔE_{intra} and the atomic

341 multipoles) in a polarizable continuum model (PCM) with a dielectric constant typical for organic

342 crystals.[64]

343 A conceptually easier approach is to calculate the relative lattice energies of a set of crystal

344 structures directly from periodic electronic structure methods, denoted Ψ_{crys}. The cost of this final

345 stage in a CSP_0 scales with the number of crystal structures, as well as the number of atoms and

346 electrons in the unit cell. Practically, a Ψ_{crys} calculation must use an electronic structure theory

347 that is far inferior to that which can be used for a Ψ_{mol} CSP_0. Indeed, the use of Ψ_{crys} for organic

348 crystals only became worthwhile when periodic density functional calculations could be corrected

349 for the missing electron correlation terms that generate the universally attractive dispersion energy

350 (DFT-D methods).

351 CSP_0 has provided a strong impetus for the development of diverse methods of evaluating the

352 relative lattice energies of organic crystals.[65-67] However, the computational effort required to

353 calculate the relative lattice energies of large numbers of CSP_0 structures can be wasted because

354 some of the static lattice energy minima are artifacts of the neglect of molecular motion. The

355 barriers between different minima can be lower than the zero-point energy of the molecular

356 vibrations,[68] and many lattice energy minima are not stable and may melt or merge to the same

357 free energy minimum.[69] To reduce the extent of over prediction associated with these features of

358 the lattice energy landscape, we have recently devised a workflow that leverages Molecular

359 Dynamics (MD) simulations and enhanced sampling methods to reduce a CSP_0 landscape to

18

360   those structures that are (computationally) distinct at ambient temperatures[70, 71] or rationalize that

361   only an amorphous phase is likely to form.[72]

362   3.1.2   Illustration of CSP_0 and landscape reduction for olanzapine

| Molecular diagram | Conformational variables | Orientation vectors |
| Conformer A | Conformer B | Dispersion-bound dimer |

363   Figure 3. Top row: Molecular diagram of olanzapine, the conformational variables (ϕ_1 and ϕ_2)

364   used in the CSP_0 search and structural fingerprints, and the vectors (θ_1 and θ_2) used to define the

365   relative orientation of molecules in the structural fingerprints. Bottom row: The two types of

366   conformer, and the van der Waals bound dimer of conformer A found in most observed crystal

367   structures containing neutral olanzapine (the notable exception is form IV).

368   Olanzapine (Figure 3) is a good illustration of how CSP and molecular modeling can be used as

369   a complementary technique in understanding the experimental solid form landscape of a drug. The

370   extensive work at Lilly, who developed olanzapine as a solid oral dosage form, established[29] that

371   the marketed form I is the most thermodynamically stable form, and that the metastable forms II

372 and III are always found concomitantly, in varying proportions, in microcrystalline samples

373 obtained by desolvating solvates. An academic joint CSP and experimental screening program[73]

374 characterized 56 solvates, and was able to solve the structure of form II from a suitable single

375 crystal grown by sublimation along with form III. This enabled the contribution of form II to the

376 diffraction of polycrystalline samples to be (computationally) removed to provide a powder X-ray

377 diffraction (pXRD) pattern for form III. The CSP study proposed a model for form III (denoted

378 III*), which was based on the same sheets as form II but stacked in a different fashion. The

379 existence of such similar CSP-generated structures helped rationalize the concomitancy of

380 olanzapine II and III, and for other molecules suggests the possibility of disorder[74] or polymorphic

381 domains.[75]

382   The CSP_0 landscape (Figure 4) showed that there were thermodynamically competitive crystal

383 structures that did not contain olanzapine in the dispersion bound dimer (Figure 3), despite this

384 motif being observed in all previously characterized crystal structures. Only recently have studies

385 of the growth of an olanzapine solvate from solution shown why the dimers appear to be the growth

386 unit as often assumed.[76-78] Decades after Zyprexa was developed, a new pXRD pattern was

387 observed for olanzapine produced by heat-induced crystallization from an amorphous

388 polyvinylpyrrolidone (PVP) based molecular dispersion.[79] The comparison of the pXRD pattern

389 with those simulated from the CSP_0 landscape lead to the structural characterization of form IV,

390 which did not contain the usual dimer. This is an example of the late-appearance of a polymorph

391 from heterogeneous nucleation, either through novel crystallization experiments such as attempted

392 crystallization or through changes in impurity profiles.
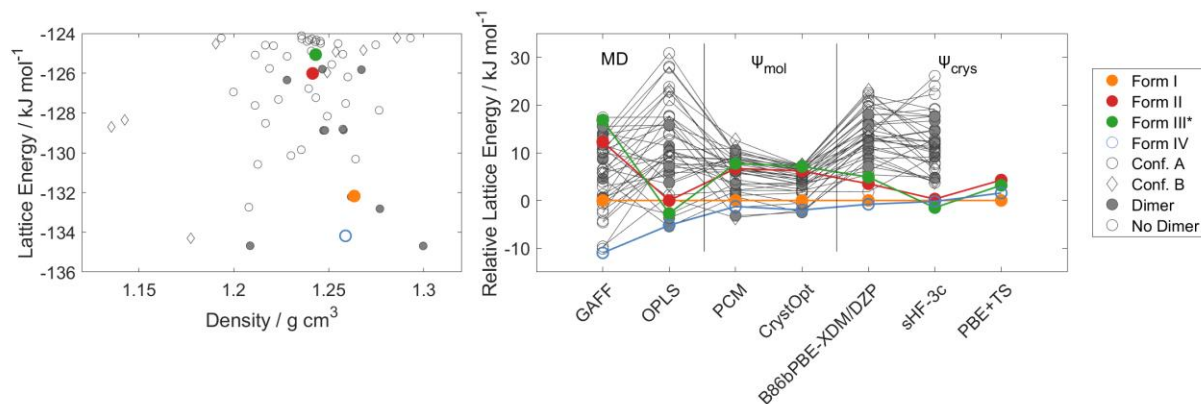
Figure 4. Left: the lattice energy landscape obtained with a Ψ_{mol} CSP study of olanzapine using CrystalOptimizer (CrystOpt) for the final lattice energy refinement, with each point representing the lattice energy and density of a mechanically stable CSP_0-generated structure. Points are classified by packing motif and those that match forms I, II and IV and the proposed model form III* given in color. Right: the comparison of lattice energies obtained from different methods. The central Ψ_{mol} contrasts the CrystOpt energies shown in the lefthand landscape with those recalculated with the intramolecular energy penalty and distributed multipoles calculated in a PCM with a dielectric constant ε=3 typical of organic crystals, as published in ref [73]. The Ψ_{crys} methods for all structures are from ref [80], and can be compared with the PBE-TS results for the experimental forms. The two MD force fields, the General Amber Force Field (GAFF) and the Optimized Potentials for Liquid Simulations (OPLS), are those used in the landscape reduction (Section 3.1.2). Conformers and the dimer are shown in Figure 3

The relative lattice energies of known and CSP-generated structures of olanzapine are typically sensitive to the computational model used (Figure 4). Reranking[80] these CSP_0 structures by single point plane-wave B86bPBE-XDM Ψ_{crys} calculations resulted in the four known polymorphs being the most stable structures, all within 5 kJ mol^{-1} of the most stable form I, and with form III* the least stable. A similar stability order is given by optimization of the Ψ_{mol} minima with periodic

411 PBE-TS, a Ψ_{crys} model often used in CSP_0, confirming that forms I and IV remain, within the

412 margin of error, comparable in lattice energy and more stable than forms II and III* (Figure 4).

413 Experimentally, form I appears to be the most stable form, but the relationship between the

414 metastable forms is complex,[81] with the free energy differences depending on temperature (see

415 3.2.1).

416   The CSP_0 lattice energy landscape of olanzapine has been reduced using our workflow[71] to

417 those structures that appear stable at ambient conditions (300 K, 10^5 Pa). The structural fingerprints

418 used to compare the different dynamic structures in supercells containing between 216 and 360

419 molecules required consideration of two torsion angles and the relative intermolecular orientation

420 of the vectors shown in Figure 3, as well as the center of mass distances. In MD simulations of

421 large supercells it is impractical and computationally demanding to use the same intermolecular

422 force field used in the ranking of CSP_0 structures. As such, a typical choice is to employ small-

423 molecule point charge classical force fields such as GAFF or OPLS. For olanzapine, the choice

424 of GAFF or OPLS makes a significant difference to the relative lattice energies of the structures

425 (Figure 4), with the latter conserving all the known polymorph structures after reduction (Figure

426 5), and providing a better agreement of their lattice energy ranking with the more accurate lattice

427 energy models used in CSP_0 (Figure 4). The very high relative energy for form III* and

428 difference from form II with GAFF may explain why it melts in the GAFF simulation but does not

429 melt when simulated with OPLS (Figure 5). There is some clustering of the structures, but not a

430 drastic reduction (45 CSP_0 structures give 35 clusters with GAFF and 38 with OPLS). However,

431 using well-tempered metadynamics (WTmetaD) to enhance crystal cell fluctuations removes

432 significantly more structures, (28 retained with GAFF, 32 with OPLS). At this stage, we observe

433   a higher proportion of the dimer-based structures than the non-dimeric structures transforming to

434   more stable or disordered structures.
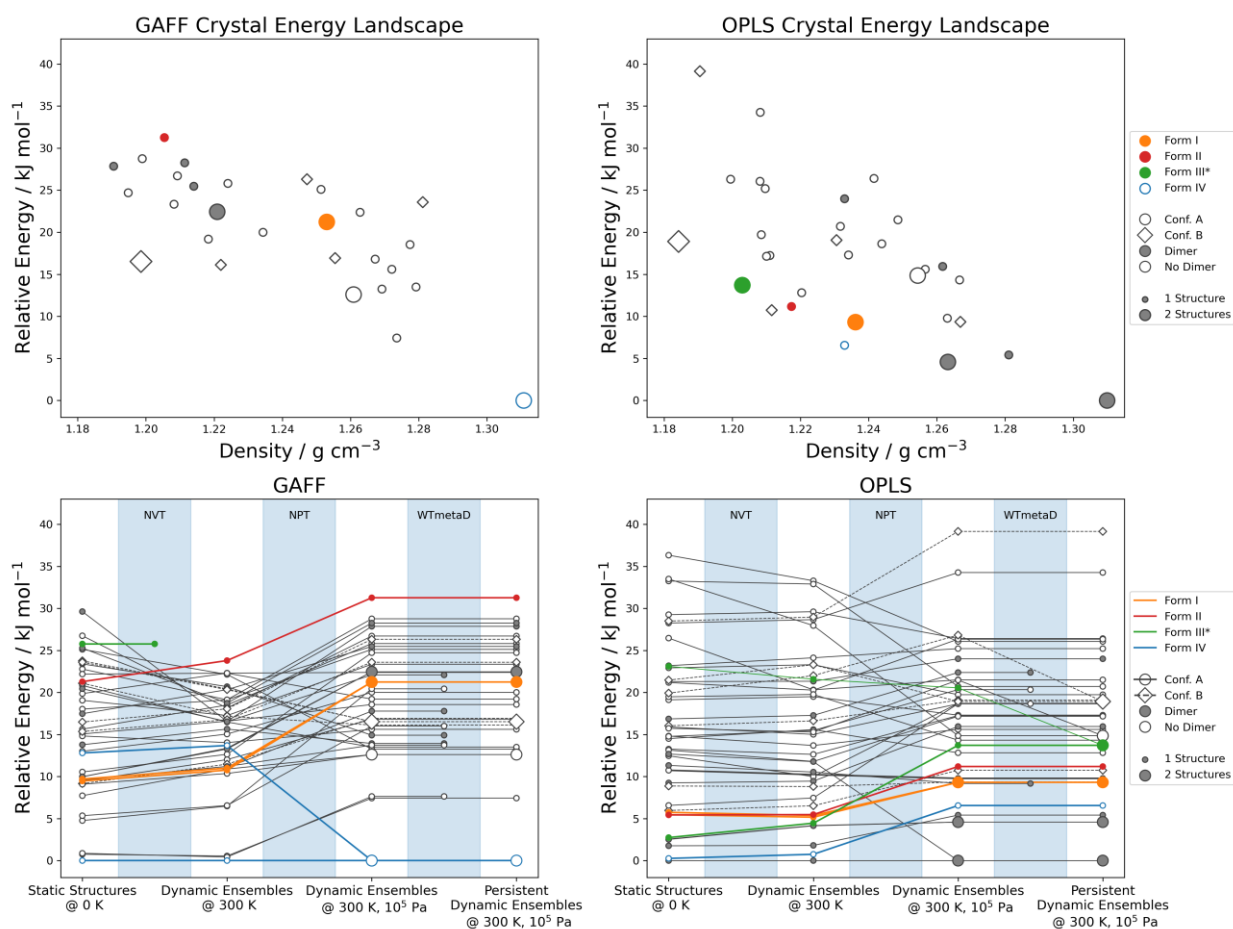
435

436   Figure 5. Top: The reduced energy landscapes of olanzapine: calculated with (left) GAFF and

437   (right) OPLS, with the size of the symbol representing the number of CSP_0 structures that have

438   merged to that form (the relative energy is the average potential energy of the structure in the MD

439   simulation at 300 K as distinct from the lattice energy (Figure 4) and so has a different energy

440   ordering). Bottom: the loss of structures (e.g. form III* through melting during the MD step)

441   resulting from the use of (left) GAFF and (right) OPLS, through the clustering and WTmetaD

442   steps. Here, relative energy of the static structures at 0 K indicates the relative lattice energy

443   computed after energy minimization, while at 300 K it is the relative potential energy averaged

444   over the last 200 ps. In the cases where relative energy at 300 K changes between the MD and

445   WTmetaD steps the initial crystal structure has transitioned to a different configuration.

446   Thus, the reduction of the energy landscape for olanzapine is not as extensive as for urea,

447   succinic acid[71] or ibuprofen.[70] While noting that the reduction extent can indeed be system-

448   dependent, we also note that in this case we had pre-selected a subset of low-energy structures that

449   seemed likely to be long lived. Most dimeric and non-dimeric structures are stable, and so the

450   explanation for non-dimeric form IV being so elusive, despite its apparent stability in lattice

451   energy, is presumably that dimers form early in the crystallization process without the presence of

452   polymer. However, the short timescale of the MD simulations in the reduction workflow means

453   that this cannot be used to estimate the free energies of the surviving structures corresponding to

454   the force field.

455   A CSP was performed on succinic acid, following the serendipitous discovery of the γ

456   polymorph, which adopts a bent conformation that is dominant in solution.[82] This structure was

457   readily predicted by a Ψ_{mol} CSP_0 search as intermediate in stability between the low temperature

458   stable β form and the high temperature α form, both of which have a planar molecular

459   conformation. The relative lattice energy is sensitive to the Ψ_{crys} dispersion correction and exact

460   crystal structure used, but all calculations with the PBE functional have the γ form as more stable

461   in lattice energy than β. Hence a polymorph that has only been observed once, despite efforts to

462   recreate the necessary crystallization conditions, would have been predicted as the most stable

463   form by a PBE-D Ψ_{crys} CSP_0 study.

464   The workflow for CSP_0 landscape reduction was developed using succinic acid as the example

465   with conformational flexibility.[71] With GAFF, a set of over 100 CSP_0 structures was reduced to

466 27 low energy, persistent crystal structures, and identified the types of disorder and stacking faults

467 that probably occur in real crystal structures, particularly the high temperature α form.

468 The reduction of a CSP_0 landscape, using the observation that the number of CSP structures is

469 reduced by an MD shake-up[83] and that hypothetical structures will melt,[84] is a major step forward

470 to reduce the number of structures that need to be considered for more accurate calculations. Other

471 methods are emerging that can use an estimate of the energy barriers between the different forms.[85]

472 3.2   Free energies of polymorphs

473 The relative stability of many polymorphs can change with temperature. In fact, a rigid-

474 molecule harmonic estimate of the free energy differences between 475 pairs of polymorphs

475 indicated the possibility of an enantiotropic phase transition in 21% of the studied systems.[86]

476 Hence it is important to go beyond 0 K lattice energies and calculate relative free energies

477 (CSD_thd). These calculations can have a considerable cancellation of errors, particularly between

478 structures with similar types of intermolecular interactions and density. However, even at 0 K the

479 zero-point motions of the atoms within the crystal can differ between polymorphs. Polymorph

480 free energies deviate from the 0 K energy differences because of different vibrational modes and

481 frequencies and different degrees of vibrational anharmonicity. These differences are apparent in

482 the IR, Raman, and terahertz spectra. In particular, the higher frequencies associated with

483 hydrogen bonding often differ between polymorphs, and there is considerable mixing of the

484 intermolecular and intramolecular modes for all but the most rigid molecules.[87] Some polymorphs

485 may even accommodate a degree of dynamic disorder under ambient conditions, e.g. torsion

486 modes may become hindered rotors for some functional groups.

487 3.2.1   *Ab initio* free energies in the harmonic approximation and beyond

488   The harmonic approximation can be used in periodic density functional codes with an increasing

489   range of density functionals and dispersion interactions[88] to determine the phonons and hence free

490   energies. As different polymorphs usually have very different unit cells, care has to be taken to

491   use a sufficiently large supercell for the comparisons (i.e. converge the Brillouin zone). The choice

492   of supercell was straightforward for the single crystal to single crystal transformations in

493   desloratadine, as the unit cells were closely related, and the harmonic phonons gave a reasonable

494   estimate of the transitions.[89] For succinic acid, calculating the Helmholtz free energy using

495   harmonic phonons stabilizes β such that β is more stable than γ at ambient temperatures, although

496   the energy differences are very small.[82] Coronene is another example where a harmonic phonon

497   calculation has been sufficient to show that the global minimum structure in a CSP_0 is actually

498   metastable at ambient temperature.[90]

499   Recently, a limited CSP_thd study of olanzapine, using embedded fragment quantum

500   mechanical methods,[91] compared the calculated frequencies of the two most stable structures (form

501   I and II) with experiment and confirmed that form I was monotropically more stable than form II,

502   with a Gibbs free energy difference that increased with temperature. An alternative electrical

503   embedding method using both DFT and single point MP2 calculations[92] gave the same stability

504   order for lattice energies as the PBE-TS calculations in Figure 4, but showed that forms IV and III

505   swapped stability order around 200 K.

506   Fundamentally, the harmonic approximation is valid only at low temperatures. It neglects

507   thermal expansion and the effects of zero-point motion on the cell parameters. Quasi-harmonic

508   calculations on carbamazepine form III[93] estimate that this introduces an error of 1-2 kJ mol^{-1} into

509   the enthalpy and entropy at ambient, but these contributions largely cancel. Hierarchical schemes

510   may allow the quasi-harmonic approximation to be applied more affordably and accurately to

511   organic crystals.[94] Several authors have included anharmonic effects and the contribution of

512   entropy at elevated temperatures for accurate estimation of free energies.[95, 96]

513   A recently demonstrated framework for the thermodynamics of polymorphs[97] uses rigorous *ab*

514   *initio* Gibbs free energy calculations, based on the streamlined development of machine-learning

515   potentials to accurately reproduce *ab initio* (PBE0-MBD) potential energy surfaces and their

516   integration with path integral methods, accounting for the quantum statistical mechanics of the

517   nuclei. The quantum nuclei corrections contribute about 1 kJ mol^{-1} to the free energy difference

518   between α and β succinic acid, correctly predicting that α is the high temperature form.[97]

519   3.2.2   Free energies from biased Molecular Dynamics

520   Molecular Dynamics (MD) when combined with techniques from statistical mechanics, can also

521   compute free energies of crystal structures. The free energy in the canonical ensemble is:[98]

522   Eq. 1

$$A = -k_B T \ \ln(Q)$$

523   where A is the Helmholtz free energy of the system, k_B is the Boltzmann constant, T is the

524   absolute temperature, and Q is the partition function, *i.e.* the Boltzmann-weighted sum over all

525   possible configurations of the system. Although force fields continue to limit accuracy, there are

526   now several methods that can compute relative free energy differences for large fully flexible

527   molecules and fully anharmonic crystals with extremely high precision, i.e. with minimal statistical

528   sampling errors. We discuss some of these methods in the following subsections.

529   From the definition of the canonical free energy as a function of the partition function, Q, it

530   follows that the free energy difference between two putative polymorphs, can be directly computed

531   from the ratio of their configurational integrals. In turn, the ratio of configurational integrals can

532   be estimated from an MD trajectory sampling the interconversion between putative polymorphs,

533   in the ergodic limit.

534   In practice, however, ergodic sampling of polymorph transitions cannot be obtained from

535   standard MD within computationally accessible timescales. As such, the calculation of free energy

536   differences between polymorphs *via* MD requires augmentation by suitable enhanced sampling

537   methods.[99, 100] A class of enhanced sampling methods that has been often used to estimate the

538   relative finite-temperature thermodynamic stability of crystal polymorphs is based on the

539   introduction of a perturbation to the Hamiltonian of the system in the form of a bias potential to

540   overcome high free energy barriers and estimate free energy differences between long-lived

541   metastable states. Methods belonging to this family include static biasing approaches, such as

542   Umbrella Sampling (US),[101, 102] as well as history-dependent biasing methods, such as

543   Metadynamics (MetaD).[103-105] Both static and history-dependent biasing methods are based on the

544   ability to sample reversibly a pathway between different crystal forms, and on the ability to

545   represent such a pathway in a low-dimensional set of collective variables (CVs). Typically, static

546   approaches, such as US, require some *a priori* knowledge of the transformation pathway between

547   crystal forms. In US, the perturbative bias introduced is harmonic in CV space, and aims to achieve

548   an exhaustive sampling of the ensemble of configurations that are projected in CV space near the

549   center of the "umbrella"-like biasing potential. Appropriately combining series of independently

550   biased simulations that partly overlap in CV space and connect the polymorphic metastable states

551   of interest, enables the calculation of global free energy surfaces and, in turn, free energy

552   differences.[101, 102, 106, 107] Unlike static approaches, history-dependent methods, such as MetaD,

553   can explore low dimensional CV spaces to autonomously identify transition pathways. In MetaD,

554   the bias potential is instead constructed adaptively as a sum of repulsive Gaussians, a function of

555   the system's position in CV space. Such a potential discourages the visitation of configurations

556   that have already been sampled, leading to an autonomous exploration of unseen configurations

557   and to the discovery of transition pathways connecting metastable states. In this case, given the

558   total bias adaptively constructed as a simulation progresses, the unbiased probability, and thus the

559   free energy of relevant metastable states, can be computed.[108-112]

560   Another approach to enhancing the sampling of polymorphic transformation *via* direct MD

561   sampling builds on the Adiabatic Free Energy Dynamics (AFED) methods developed by Rosso

562   and Tuckerman.[113] The method implemented most recently, named crystal-AFED,[114] is a variant

563   of the driven-AFED algorithm,[115-117] where CVs, such as the supercell parameters or order

564   parameters capturing the local arrangement of molecules in the crystal, are adiabatically decoupled

565   from fast degrees of freedom and sampled at an artificial high temperature. In this case the bias

566   introduced by the sampling algorithm can be rigorously removed by applying appropriate

567   reweighing techniques.[112]

568   In all methods discussed, the sampling efficiency and the accuracy of the free energy estimates

569   depend on the ability of CVs to resolve degeneracies between metastable states of interest and to

570   capture the slow transition modes associated with a polymorphic phase transformation. As such,

571   for most molecular systems, the calculation of free energy differences between crystal forms

572   requires the identification or the development of *ad hoc* CVs, and the ability to efficiently compute

573   such CVs on-the-fly as a dynamic simulation progresses.[118-124] These aspects become increasingly

574   problematic with the complexity of the growth units considered,[125] and this directly impacts the

575   applicability of MD-based sampling approaches to large scale CSP_thd studies, where relative free

576   energy differences between hundreds of crystal forms, often exhibiting radically different local

577   structure, must be computed. As such, while direct sampling methods can yield a rich insight into

578   the mechanism of phase transitions and can in principle be used to investigate nucleation of

579   competing polymorphs in realistic environments, their application to CSP_thd is still somewhat

580   limited.

581   ### 3.2.3   Free energies by Einstein Crystal

582   To circumvent the need to directly sample phase transformations, a common strategy is that of

583   taking advantage of thermodynamic cycles connecting the metastable state to a suitably defined

584   reference state. There are many literature examples where a reference state is chosen for which

585   free energy can be computed analytically or numerically.[126-128] In the case of fluid phases, the most

586   suitable reference state is an ideal gas, and for the solid phases, the appropriate reference state will

587   be an Einstein crystal, introduced in 1984 by Frenkel and Ladd. This is a rigorous method to

588   compute the absolute free energy of solids.[126] An Einstein crystal is structurally identical to the

589   crystal of interest, with no interparticular/intermolecular interactions.[126] Each atom in the Einstein

590   crystal is attached to its lattice site *via* harmonic springs. Once the reference state's free energy is

591   known, the reference state is connected to the system of interest *via* a reversible path. Free energy

592   differences are computed along the reversible path as a ratio of partition functions.[129] The most

593   common methods to compute free energy differences are the free energy perturbation (FEP)

594   method and thermodynamic integration (TI).[130] The integration path should be free from phase

595   transitions.

596   Vega proposed a variant of the Einstein crystal method known as the Einstein molecule

597   method.[127] The key difference between these two methods lies in the implementation. In the

598   Einstein crystal method, the system's center of mass is held constant to avoid divergence at low

599   spring strength. In the Einstein molecule method, one of the atoms in the simulation box is held

600   constant to achieve the same.[127] These methods start from an Einstein crystal/molecule, then turn

601   on the interparticular/intermolecular interactions using the perturbation method. Finally, harmonic

602 springs attached to the lattice sites are removed to recover the solid of interest using TI. The

603 methods based on Einstein crystal and Einstein molecule references are often collectively called

604 "Frenkel-Ladd" methods, and they have been extensively used to compute the free energy of

605 atomic solids and molecular solids. Versions of these methods have also been implemented in

606 MD packages.[131-134]

607 The Frenkel-Ladd methods obtain polymorph free energy differences by subtracting the absolute

608 free energies evaluated for two different polymorphs. Typically, the free energy difference is small

609 (a few kJ mol^{-1}) compared to the absolute values. Hence to obtain a meaningful small free energy

610 difference, two large free energy differences (from Einstein reference to each polymorph) must be

611 computed with high precision. The lattice switch methods provide a more direct polymorph-to-

612 polymorph pathway. Lattice Switch Monte Carlo (LSMC)[135] uses an energy gap order parameter.

613 The energy gap is obtained for any configuration by one-to-one mapping the thermal

614 displacements from the 0 K structure of one polymorph onto the 0 K structure of the other

615 polymorph. Umbrella sampling or other strategies can then drive the energy gap from negative

616 values through gateway states, where the energy gap is zero, to positive values for a reversible

617 pathway from one polymorph to the other. The recipe, in principle, works for arbitrarily complex

618 crystal structures, although LSMC itself has mainly been used in high accuracy calculations for

619 atomic solids.[135-137] A diabat free energy variant of LSMC makes further use of the exact Zwanzig-

620 Bennet relationship[138, 139] between free energy diabats.[140] This new method shows promising

621 precision and efficiency for polymorphs of atomic systems,[141] as well as molecular systems[142]

622 (requiring only two unbiased MD simulations in some cases). Our application to carbamazepine

623 demonstrated precision at a level ±0.01 kcal mol^{-1} (±0.04 kJ mol^{-1}) in computed free energy

624 differences with just 5 ns of computing time per polymorph pair.[142]

3.3   Open problems in CSP and free energy calculations

626   There are many open problems in the study of polymorphism.[143] The ideal polymorphs are those

627   that are predicted to be the most stable under any condition of temperature, pressure, or relative

628   humidity to which the drug product could be exposed. However, this is not generally the case for

629   many systems and the experimental observation of phase transformations, particularly in the solid

630   state, can be difficult.[144] Indeed, the prediction of free energies is particularly critical when there

631   is a major kinetic barrier to the transformation between structures, such that the transformation is

632   not readily or easily observed experimentally.

633   The possibility of a change in stability order of polymorphs is restricted to cases where the lattice

634   energy difference is comparable to the relative thermal contributions to free energy. However, as

635   shown clearly in Figure 4, the error in the relative lattice energies is still substantial, particularly

636   with the potential energy surfaces (force fields) that can be used to simulate anharmonic effects,

637   i.e., realistic molecular movement, including thermal expansion.

638   In choosing a method for computing polymorph free energies, several factors must be

639   considered. These include issues of both accuracy and precision. Accuracy pertains to the model

640   chemistry, e.g. an *ab initio* description or an empirical force field both of which include numerous

641   options with varying accuracy. Precision pertains to the ability to sample completely and

642   efficiently without assuming harmonic vibrations or other approximations beyond those inherent

643   to the model chemistry. Considerations of accuracy and precision are typically interrelated. For

644   example, high level electronic structure calculations offer better accuracy than most force fields,

645   but they are limited to harmonic approximations. In contrast, force fields enable calculations at

646   the 0.01 kJ mol^{-1} level of precision, but the predictions are often hampered by questions about

647   force field accuracy.

648   The comparison of a CSP output with the known crystal forms usually poses the question as to

649   which of the unknown structures could be observed as relevant polymorphs. This is where CSP

650   can play a significant role as a complement to solid form screening.[45] There are often "predicted"

651   structures that are competitive with, or even more stable than the known forms,[145] raising the

652   question as to whether knowing the crystal structure of the targeted form can suggest an experiment

653   to produce the first sample. When the targeted polymorph is denser than the observed polymorphs,

654   further calculations of the lattice enthalpy as a function of the applied pressure may suggest that

655   crystallization under pressure would favor the targeted form.[146, 147] The calculation of energy-

656   structure-function maps,[148] where the properties for each CSP-generated structure are

657   calculated,[149] has been powerful in the search for novel functional materials. It can also be used

658   for testing whether changes in experimental conditions are likely to be useful in polymorph

659   discovery, for example, the calculation of the anisotropy of the diamagnetic susceptibility tensor

660   of different polymorphs can indicate whether the polymorphic outcome may change if the

661   crystallization occurs within a magnetic field.[90] The most direct method to target a specific crystal

662   structure is when there is another molecule which adopts the desired crystal structure and can be

663   used as a template in a crystallization experiment.[150] For example, the first catemeric polymorph

664   of carbamazepine (form V) and form III of cyheptamide were found[151] by sublimation onto a

665   crystal of dihydrocarbamazepine form II, and tolfenamic acid forms VI and VII were discovered

666   by sublimation onto mefenamic acid form I and a tolfenamic acid:flufenamic solid solution,

667   respectively.[152]

668   4   Solubility Determination

669   Critical to the development of a crystallization process is the determination of the equilibrium

670   solubility of the given compound in the process solvent(s). Solubility is the concentration limit of

671 a chemical compound where the chemical potentials of the solid and solution phases are equal.

672 Supersaturation, the concentration of the solute above the saturation concentration of the specific

673 form, provides the driving force for crystallization. While it is generally straightforward to measure

674 the solubility of a stable polymorph, metastable polymorphs, which by definition create

675 supersaturated solutions, tend to transform in solution during the experiment into a more stable

676 form.[153] Computational methods to predict phase equilibria and free energy differences between

677 solid and fluid phases are therefore critical for studies of polymorph-specific crystal nucleation

678 and growth, and have great potential to aid solubility estimation tasks during crystallization

679 process development of API molecules. Solubility can either be calculated from considering the

680 coexistence of the two phases (Section 4.1) or by a thermodynamic cycle (Section 4.2).

681 4.1   Methods based on coexistence

682   Direct coexistence simulations bypass the need for free energy calculations and directly provide

683 an estimate of the solubility limit. Basically, a long simulation with the crystalline material and

684 the solution in contact (usually at isobaric and isothermal conditions) allows the two phases to

685 reach equilibrium with each other. Direct coexistence simulations are simple and useful for initial

686 solubility estimates, but they have two important limitations.[154, 155] These simulations are prone

687 to errors arising from long time scale, activated processes, such as attachment and detachment at

688 kinks,[156, 157] 1D nucleation of kinks,[158-160] and 2D island or pit nucleation.[154, 161-165] Simulations

689 with special "everkinked" crystal orientations can eliminate the row/kink and island/pit nucleation

690 phenomena,[162, 166] but the attachment and detachment rates at kink sites still set a fundamental limit

691 on the efficiency of direct coexistence results. These limitations can be effectively mitigated by

692 using enhanced sampling methods to facilitate an exhaustive sampling of the attachment and

693 detachment processes to and from kink sites, as recently demonstrated by Bjelobrk et al. for

694   organic molecules[167] and salts in solvent-antisolvent mixtures.[168] However, direct coexistence

695   simulations can only estimate equilibrium concentrations, e.g., the equilibrium solubility limit, the

696   melting temperature, or the partial pressure for evaporation. They cannot estimate the chemical

697   potential differences at non-equilibrium supersaturated conditions that drive crystallization.

698   4.2   Methods based on thermodynamic integration

699   An alternative route is to compute the dissolution free energy in stages *via* a thermodynamic

700   cycle of subliming the crystal into the gas phase and then solvating the molecule, an approach

701   shown adapted to thermodynamic integration (TI) in Figure 6. We have already seen (Section

702   3.2.3) how Frenkel-Ladd methods (TI for solids) starting from the Einstein crystal or Einstein

703   molecule can compute free energies of solid phases.[169, 170] Starting from a suitable reference model

704   for the gas phase molecule, TI methods can also provide the free energy of molecules in solution

705   as a function of temperature, solute concentration, and pressure. The typical gas phase reference

706   for small and relatively rigid molecules is the free energy with harmonic vibrations. Figure 6

707   shows the alternative centroid reference of Khanna *et al.*[169] The centroid is comprised of all atoms

708   in the solute molecule, tethered *via* springs to their collective center of mass. It has the same mass

709   and the same number of vibrational and rotational modes as the final molecule. The centroid is

710   converted to a fully anharmonic molecule by a TI that turns on all intramolecular interactions such

711   as bonds, angles, dihedrals and pairwise interactions within the molecule, e.g. Lennard-Jones and

712   coulombic interactions.[169] This one extra step, the centroid-to-molecule transformation, provides

713   the absolute free energy for fully anharmonic and multiconformer gas phase molecules. The next

714   step, whether starting from a harmonic reference or centroid reference, is to compute the free

715   energy of solvation. Note that the solute free energy in the gas phase is a function of its partial

716   pressure, and that in the solution phase it is a function of its mole fraction.
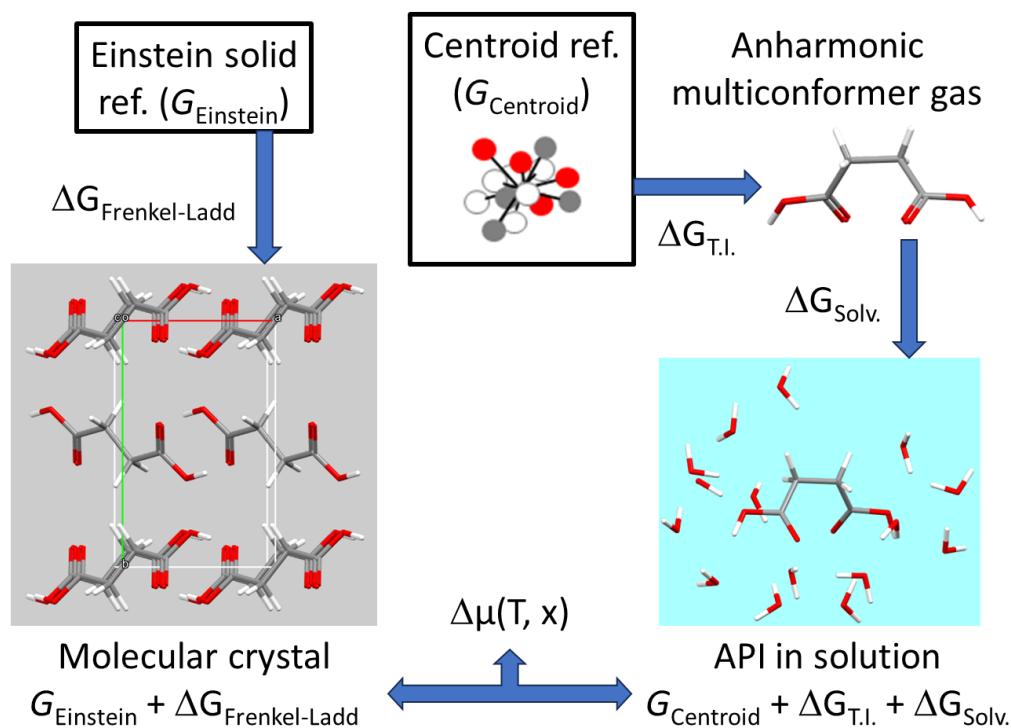
Figure 6. Thermodynamic integration starting from two reference models to obtain the free energy of the crystal and that for the molecules in solution. The calculations are illustrated for the β form of succinic acid, which has a planar conformation in the crystal but the dominant gas and solution phase conformation is bent.[82]

The solubility can be computed by equating the chemical potential of the compound in the crystal phase with that in the solution phase, i.e. $\mu_{crystal}(T) = \mu_{solution}(T, x)$ where x is the mole fraction of solute in solution. The equality of chemical potentials is satisfied for the special composition x_{sat}, i.e. the solubility limit.

4.2.1   Heat of sublimation

The heat of sublimation, the energy difference between the crystal and gas phases, is usually the best experimental test of the energy scale of the intermolecular interactions within the crystal and hence the lattice energy. Computationally, we can estimate the heat of sublimation from the slope

730 of Clausius-Clapeyron plots, which require vapor pressure data at various temperatures.[170] Vapor

731 pressures at each temperature are computed by equating the chemical potential of the crystal phase

732 with that of the gas phase.[170] Crystal and gas phase free energies are computed using the methods

733 mentioned in Sections 3.2.3 and 4.2.

734 4.2.2   Solvation energies

735   For the solution phase, we add the solvation free energies to the molecule's intramolecular and

736 translational free energy. Solvation free energy refers to the work required to insert one solute

737 molecule into the given solvent. Like Widom's method,[171] direct insertion of solute molecules with

738 complex molecular structures into the highly dense solvent phase will be difficult. Alternatively,

739 several authors have come up with new strategies based on Monte Carlo methods to address this

740 problem. For example, methods like Configuration Bias Monte Carlo methods[172] (CBMC) and

741 Continuous Fractional Monte Carlo methods[173] (CFMC) will improve the statistics for the

742 insertion of solute molecules in the solvent.
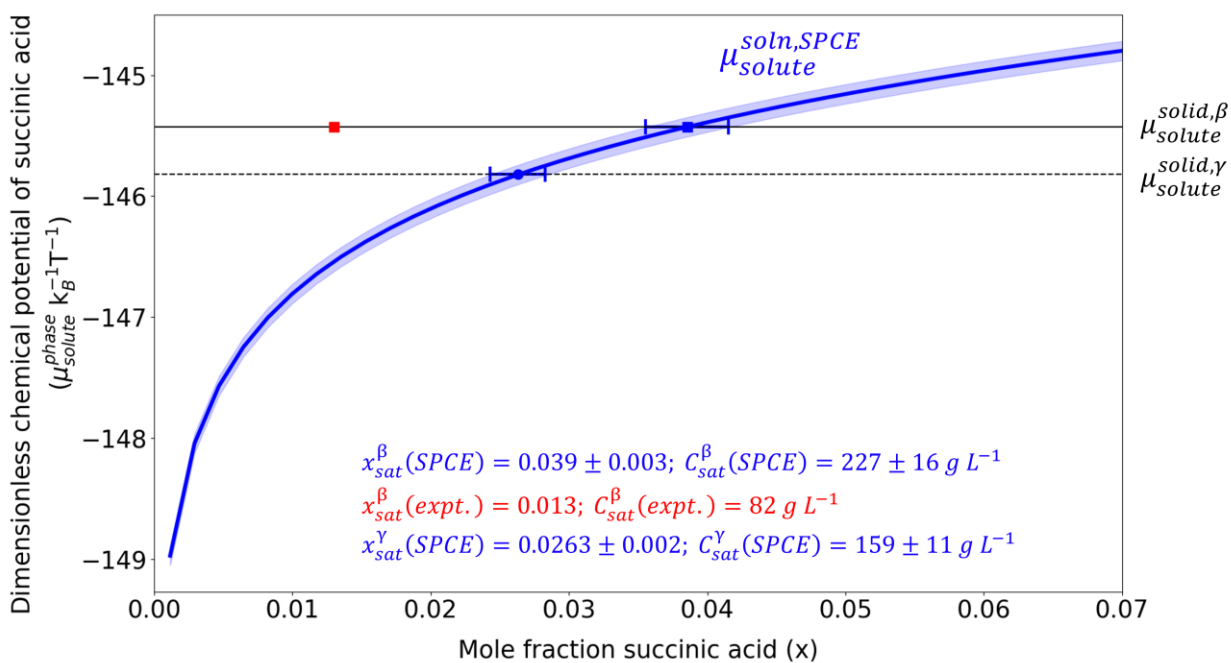
743   Other methods combine MD simulations along with FEP or TI to compute solvation free energy.

744 Most of these methods use an external repulsive potential (Weeks-Chandler-Anderson (WCA)

745 potential[174] or Born potential) to create a cavity for solute insertion. Shivakumar *et al.* developed

746 a method that uses MD with an FEP scheme to compute absolute solvation free energies of small

747 neutral molecules in various solvents.[175] Mobley and Guthrie computed the free energy of

748 hydration for nearly 650 molecules and constructed a database of solvation free energies for

749 explicit and implicit solvent models.[176] They report an rms error of 6.3±0.3 kJ mol^{-1} between the

750 computed and experimental hydration free energies using GAFF. Li *et al.* proposed a TI based

751 scheme with MD simulations to compute the solvation free energies of sparingly soluble solutes

752 in different solvents. The strategy involves three main steps. First, the solvent is evacuated from

37

753   a cavity with an external repulsive potential, then the solute molecule is inserted into the cavity,

754   and finally the external potential is removed to recover the state of interest.

755   Recently, Khanna *et al.* proposed a decoupling approach to compute the solvation-free energies

756   of solutes in each solvent.[177] In this method, the authors invoke a shorter thermodynamic cycle by

757   decoupling the solute-solvent interactions in the solution phase. Then they gradually turn on the

758   Lennard-Jones and coulombic interactions in two steps to compute the solvation free energy in

759   each solvent. The added advantage with the recent two methods is that these methods are

760   compatible with popular open-source MD engines such as LAMMPS and GROMACS.

761   4.2.3   Pilot Compound Results – succinic acid

762   The approach reported in references [177] and [169] has been used to predict the solid-vapor and solid-

763   solution equilibria of β- and γ-succinic acid at various temperatures, along with the driving forces

764   (chemical potential difference for solute in solution phase and crystalline phase) for crystallization.

765   Shown in Figure 7 is the chemical potential computed for succinic acid as a function of solution

766   phase mole fraction at 300 K, along with the computed chemical potentials for the β and γ

767   polymorphs. The vapor pressures and solubility predictions for succinic acid when plotted against

768   inverse temperature give rise to the Clausius-Clapeyron (Figure 8) and van't Hoff (Figure 9) plots,

769   respectively.

The figure shows a plot with "Dimensionless chemical potential of succinic acid (μ_{solute}^{phase} $k_B^{-1}T^{-1}$)" on the y-axis and "Mole fraction succinic acid (x)" on the x-axis. The curve is labeled $\mu_{solute}^{soln,SPCE}$, with horizontal lines labeled $\mu_{solute}^{solid,\beta}$ and $\mu_{solute}^{solid,\gamma}$.

$x_{sat}^{\beta}(SPCE) = 0.039 \pm 0.003; \; C_{sat}^{\beta}(SPCE) = 227 \pm 16 \; g \; L^{-1}$

$x_{sat}^{\beta}(expt.) = 0.013; \; C_{sat}^{\beta}(expt.) = 82 \; g \; L^{-1}$

$x_{sat}^{\gamma}(SPCE) = 0.0263 \pm 0.002; \; C_{sat}^{\gamma}(SPCE) = 159 \pm 11 \; g \; L^{-1}$

770

Figure 7. Chemical potential as a function of solubility mole fraction in aqueous solution at 300 K and 1 atm pressure (adapted from Figure 9 in Khanna *et al.*[170], reprinted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com)), along with computed chemical potentials of the β and γ succinic acid polymorphs. The force field model used for succinic acid is GAFF, and that used for water is SPCE.[178] The computed saturation solubility, x_{sat}, of each polymorph lies at $\mu_{crystal}(T) = \mu_{solution}(T, x)$, where the chemical potential of the compound in the crystal phase and that in the solution phase intersect. The width of the curve represents plus and minus one standard deviation in the predicted chemical potential, i.e., the "error bar." The error bars reflect statistical uncertainty (i.e. precision) in the calculations, as opposed to bias errors from force field inaccuracy which are the main reason for the discrepancy with experiment. In conversion of mole fractions to C_{sat}, we assumed negligible volume of mixing, i.e. that the volume of the solution was equal to the volume of water at 300 K plus the volume of solid succinic acid at 298 K.

784
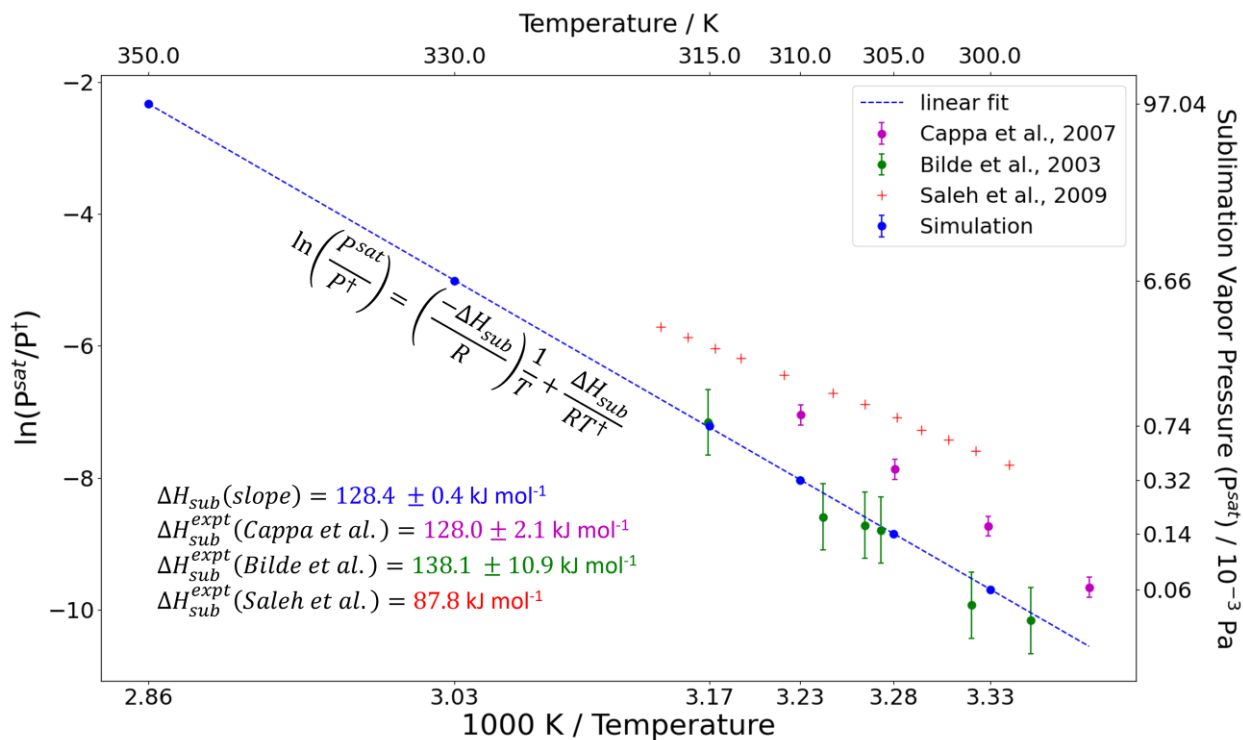
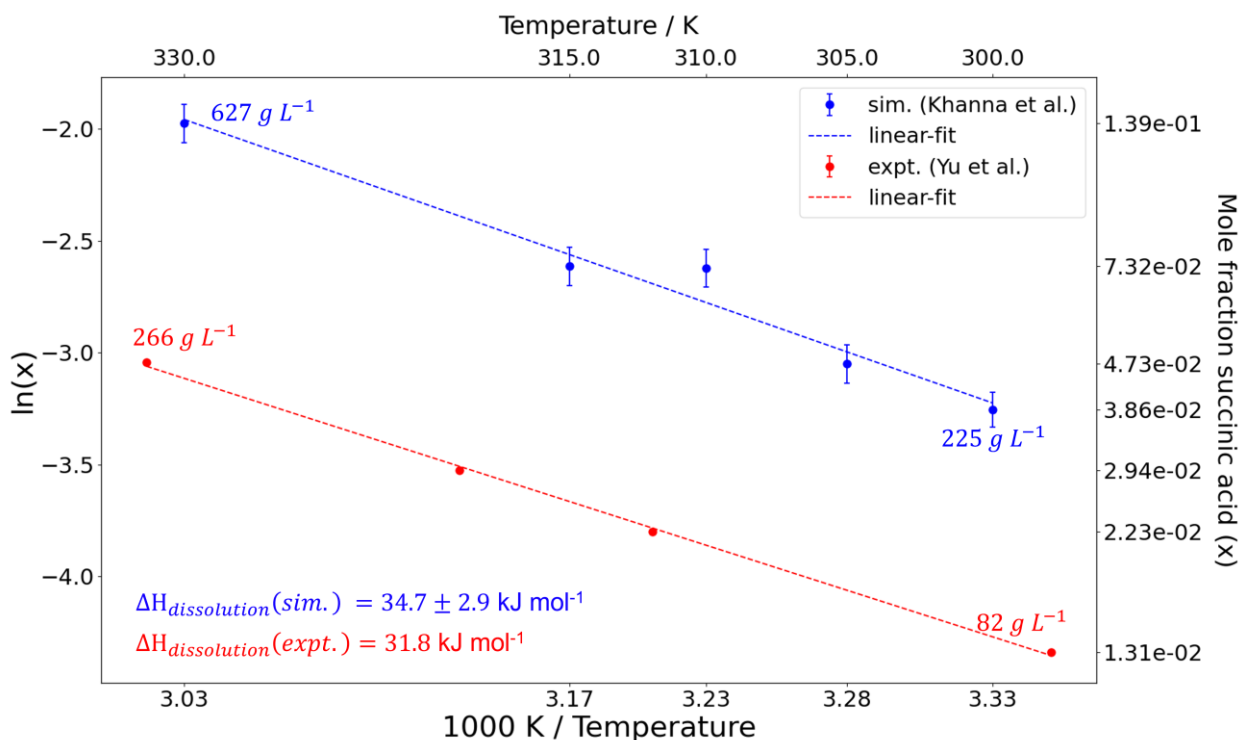Figure 8. Clausius-Clapeyron plot of β succinic acid sublimation vapor pressure and temperature, where $P^{\dagger} = 1$ Pa, and $T^{\dagger}$ is the corresponding sublimation temperature. The tick marks on the secondary ordinate axis on the right-hand side of the figure correspond to the simulation data points. Adapted from Figure 6 in Khanna et al.,[170] reprinted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com), including experimental data from Cappa et al.,[179] Bilde et al.,[180, 181] and Saleh et al.[182]

Figure 9. Van't Hoff solubility plot for β succinic acid in aqueous solution at atmospheric pressure. The tick marks on the secondary ordinate axis on the right hand side of the figure correspond to the experimental (taken from Yu *et al.*[183]) and simulation data points (taken from Khanna *et al.*[170]), with tick marks for overlapping points not shown for clarity. Adapted from Figure 11 in Khanna *et al.*,[170] reprinted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com). In conversion of mole fractions to C_{sat}, we assumed negligible volume of mixing, i.e. that the volume of the solution was equal to the volume of water at the appropriate temperature plus the volume of solid succinic acid at 298 K.

On the Clausius-Clapeyron plot (Figure 8) the simulation points lie almost perfectly on a straight line although no such requirement is imposed on the simulation algorithm (of course the experimental data are equally unaware that they are expected to follow the same trend). The slope of the line gives the sublimation enthalpy, which is in reasonable agreement with the range of

41

804   experimental values. At the current time it is not possible to say anything general about vapor

805   pressure predictions since Figure 8 is the first vapor pressure prediction for organics that we know

806   of in the literature. Nevertheless, we can conclude that the estimation of sublimation vapor

807   pressure is now possible for organic molecules and may be a useful addition to workflows for

808   digital drug product design.

809   One word of caution is worth mentioning. Vapor pressure predictions exponentially magnify

810   small errors in the chemical potential calculations (regardless of whether these errors arise from

811   simulation methodology errors or force field inaccuracies), as explained by Khanna *et al.*[169] in

812   connection with eq. 16b in that paper. In their calculations, an error of 0.05% in absolute chemical

813   potential for succinic acid becomes a 10% error in vapor pressure prediction.

814   The computed slope (enthalpy of solution) tends to be in close agreement with experiments, but

815   predictions for the intercept (related to entropy of solution and the absolute solubility) are less

816   reliable as seen in Figure 9.[184-186] One reason is that Van't Hoff plots for solubility-temperature

817   relationship exhibit non-linearities over a wide temperature range.[187] The nonlinear 1/T

818   dependence emerges because enthalpy and entropy of solvation are not truly temperature

819   independent quantities. Moreover, as seen in Figure 7, the absolute solubility (related to the

820   intercept in the Van't Hoff plot) is highly sensitive to the computed chemical potential. If the

821   curve of chemical potential of the solute in solution, $\mu_{solute}^{solution}$, is shifted up by approximately 1

822   k_BT or if the chemical potential of the pure solid solute is shifted down by approximately 1 k_BT the

823   predictions would align with experiments. This difference is less than 1% of the free energies

824   being computed (1 part in ~ 145), which demonstrates the level of accuracy required to correctly

825   predict absolute solubility. This is unlikely to be achieved very often with the current generation

826   of classical force fields, but with just one experimental data point, the simulated line can be fixed

827   in the correct position to follow the remaining experimental data.

828   Polymorph stability is demonstrated in Figure 7 where the chemical potential of the γ crystalline

829   form of succinic acid is also plotted. The resulting chemical potential for the γ form is lower than

830   for the β form (more negative), leading to a lower solubility in solution, indicating a more stable

831   solid form at 300 K and 1 atm pressure. However, the two forms differ in their chemical potential

832   by only ~ 0.4 k_BT molecule^{-1}, which is small indeed compared to the absolute scale of ~ −145 k_BT

833   molecule^{-1}. It seems reasonable to conclude that the two forms are close in stability, so close in

834   fact that it may be difficult to discriminate between them, as also shown by the relative lattice

835   (Section 3.1.2) and free energies (3.2.1). However, to do this experimentally will first require an

836   experimental protocol for producing and isolating the γ form which has only been observed once

837   in the lab as a concomitant solid with the commonly observed β form.[82]

838   Another method of calculating absolute solubility by the thermodynamic cycle is based on using

839   the Ψ_{mol} estimate of lattice energies (section 3.1.1) and the harmonic phonons (3.2.1) to evaluate

840   the heat of sublimation, along with either implicit or explicit solvation models for the heat of

841   solvation. This has been shown to rival the accuracy of informatics models for calculating absolute

842   solubility.[188] In the case of succinic acid, computing hydration free energies from atomistic MD

843   simulations using FEP methods and GAFF combined with the water SPCE force field gave a

844   hydration energy of -57.47 kJ mol^{-1}, in somewhat better agreement with experiment (-61.08 kJ

845   mol^{-1})[189] than various implicit solvation models. A Ψ_{mol} estimate of the lattice energies, and Ψ_{crys}

846   estimate of thermal corrections from harmonic phonon calculations gave a heat of sublimation of

847   121.04 kJ mol^{-1} (Fowles et al.[188] used an experimental value[190] of 123.2 kJ mol^{-1} which can be

848   compared with the values on Figure 8). Combining these estimates gave an aqueous solubility of

849   0.447 mol L^{-1}, equivalent to 52.75 g L^{-1}, which is within reasonable agreement with experimental

850   determinations ranging from 70 to 82 g L^{-1}. This was more accurate than the best machine learning

851   estimate, yet the agreement with experiment is very dependent on the quality of electronic structure

852   method and choice of empirical force field used.[188]

853   4.3   Open problems in the calculation of solubility

854   Solubility prediction is an ongoing challenge for "first principles" methods that do not require

855   any experimental input. The simplest strategy accounts for solvation through an implicit solvation

856   model, often based on Poisson-Boltzmann equations[191] or polarizable continuous medium

857   models[192] in combination with density functional theory calculations. A related strategy, used

858   extensively in crystallization, is the conductor-like screening model (COSMO). This uses *ab initio*

859   calculations to obtain "sigma profiles" (surface charge density profiles) specific to each solute and

860   solvent molecule. Solvation free energies are estimated *via* a thermodynamic cycle of melting and

861   mixing with the solvent sigma profiles This method is a conductor-like screening model based on

862   the use of solvation thermodynamics and can use the growing open-source database of "sigma

863   profiles" for different solvents.[193] Although the COSMO approach offers a route to varying both

864   single solvents and mixed solvents (especially, solvent-antisolvent mixtures) into crystal growth

865   and morphology modeling, there is the limitation of the requirement of experimental melting data

866   for the drug crystals that may not be obtainable when there is thermal degradation.

867   Moving to explicit solute-solvent simulation methods, direct coexistence simulations are

868   susceptible to undetected errors, and they cannot provide thermodynamic driving forces beyond

869   the solubility limit. Methods based on harmonic approximations can be implemented with

870   electronic structure theories that remain accurate across solutes and solvents, but harmonic

871   approximations and continuum solvent approximations become inaccurate for large molecules

872 (especially in fluid phases). Methods based on thermodynamic cycles require precise free energy

873 calculations for the crystal phase, the gas phase, and the solvation free energies. All three free

874 energy calculations can now be done precisely, but accuracy remains limited even for the best

875 available force fields. Given the difficulty in experimentally determining the thermodynamic

876 differences between polymorphs, or even their melting points and heats of fusion, further progress

877 in the calculation of solubilities is both necessary and expected. Efforts to predict solubility may

878 benefit from machine learning methods, which require large experimental datasets for training.[188]

879 An open challenge in solubility prediction is to combine machine learning with physical principles,

880 in ways that can guide the rational selection of solvents.[194]

881 5   Crystal Growth and Morphology Prediction

882 Crystal morphology (by which we mean the particular faces exposed on the crystal surface

883 together with their relative areas) plays a significant role in pharmaceutical manufacturing.

884 Needle-like and plate-like crystals are to be avoided because they are difficult to process in both

885 upstream (crystallizer-filter-dryer) and downstream (formulation and tableting) operations.

886 However, due to the large number of crystalline products in development, coupled with the vast

887 choice of growth conditions, including choice of solvent (and possibly solvent mixture and/or

888 choice of antisolvent), temperature, supersaturation, pH, and impurities/additives, it is a great

889 challenge to obtain specific crystals which have a particular product functionality and morphology

890 in an acceptable process development timescale. *In silico* methods hold promise of providing

891 useful tools to guide experiment and pre-screen the desired crystals more cheaply and efficiently

892 In many instances the goal is to produce well-faceted prism-like crystals of the desired

893 polymorph in the size range $30 - 80$ μm.[195] Such crystals are generally formed in a layer-by-layer

894 fashion. The usual transport picture is that solute molecules diffuse from the bulk solution across

895 a boundary layer to the crystal-solution interface (3D diffusion), where they adsorb on the crystal

896 terraces, then diffuse across the terraces (2D diffusion) and incorporate into vacant sites (called

897 kink sites) on the steps, thereby causing the steps to "flow across the surface" to create the

898 layers.[196, 197] At low supersaturation, the steps are generated by screw dislocations on the crystal

899 surface leading to the famous spiral growth mechanism of Burton, Cabrera and Frank,[198] or at high

900 supersaturation they correspond to the faceted edges of a 2D nucleus that has formed on the crystal

901 face.[199-203] Under such a layer-by-layer mechanism the growth rate, G_{hkl} (in units of nm s^{-1}), of

902 each face (hkl) is given by

903 Eq. 2
$$G_{hkl} = \frac{h_{hkl}}{\tau_{hkl}}$$

904 where h_{hkl} is the step height (nm), and τ_{hkl} is the time taken (s) to lay down a layer. Under normal

905 conditions for API molecules G_{hkl} is in the range 5-100 nm s^{-1} (or about 5-100 unit cells per

906 second).

907 In many industrial crystallizations, the slowest process in this picture is solute incorporation at

908 kink sites.[198, 200, 204] Mechanistic models based on this assumption are widely used in the literature.

909 The step height is a simple multiple (typically ½ or 1) of the interplanar spacing, d_{hkl}, and for spiral

910 growth τ_{hkl} is the rotation time of the spiral, which involves the step velocity, v_i^{step}, for each of the

911 i steps on the spiral, and critical length of each spiral side.[205] The critical length can be reliably

912 calculated using the Gibbs-Thomson equation[206] (also see Fig. 3 and eq. 23 in ref [202]). Step

913 velocity, however, is more challenging. The normal assumption is that the step remains stationary

914 until it has reached its critical length, whereupon it travels at constant velocity across the surface

915 of the face (this is the Voronkov assumption;[207] see Teng *et al.*[206] for an assessment of this model).

916 For the special cases of Kossel growth units (cubic "molecules" with six identical faces each

917 exposing identical "broken bonds" with neighboring molecules) and centrosymmetric growth units

918   (molecules with an inversion center such as planar succinic acid and naphthalene) step velocity is

919   calculated using

920   Eq. 3 $$v_i^{step} = a_{p,i}\rho_i u$$

921   where $a_{p,i}$ is the distance that edge i propagates with the addition of a single row of growth units,

922   ρ_i is the density of kink sites along the edge; and u is the kink rate, which represents the net

923   incorporation rate of growth units into the kink site (rate of attachment minus rate of detachment).

924   The propagation distance is determined from the size of the growth unit and crystallographic

925   considerations. At low levels of supersaturation such as often occur in cooling crystallization, the

926   kink density is taken to be constant (independent of supersaturation) and equal to its value for

927   single-height kinks at saturation conditions, given by the Boltzmann distribution[208]

928   Eq. 4 $$\rho_i = \frac{2}{2+e^{\beta\phi_i^{kink}}}$$

929   where ϕ_i^{kink} is the "kink energy" of edge i, and β takes its usual meaning of $1/k_BT$. The kink energy

930   is half the strength of total interaction energy between neighboring growth units in the kink

931   direction – it is taken to be half the total bond energy between neighboring molecules in the kink

932   direction so that each molecule has the same "broken bond" energy – and is a key parameter in the

933   theory of crystal growth.

934   For vapor (sublimation) growth, ϕ_i^{kink} is the solid state bond energy (calculated using an atom-

935   atom force field such as GAFF) in the kink direction between neighboring growth units on the

936   edge. For solution growth, this kink energy must be "solvent-modified." This is often

937   accomplished using the classical Dupre interface model in which the interfacial energy between

938   crystal and solvent at the kink site is given by the cohesive energy minus the adhesive energy.[209]

939   Solubility parameter methods are often used to capture the adhesive energy between the solvent

940   and crystal surface.[210-212] Different solvents modify the kink energy by different amounts, which

941   is how the "solvent effect" is captured in the model.

942   For centrosymmetric growth units, the kink rate is given by

943   Eq. 5 $$u = j^+ - j^-$$

944   where the attachment rate, j^+, and detachment rate, j^-, are given by the following rate

945   expressions.[157, 213]

946   Eq. 6 $$j^+ = k^+ x = k^+ x_{sat} S$$

947   Eq. 7 $$j^- = k^-$$

948   The attachment rate is first order with respect to the solute mole fraction in solution, x (which is

949   related to supersaturation, $S = x/x_{sat}$), and the detachment rate is zeroth order. The attachment rate

950   constant, k^+, is normally assumed to be site-independent (i.e., depends only on solvent, but not the

951   specific surface site). This is assuming that the complexities of attachment that affect absolute

952   growth rates (Section 5.3) are the same for all surfaces and so cancel for morphologies. The

953   detachment rate constant, k^-, depends on the bonding structure of the site (in addition to the

954   solvent). It follows that[157]

955   Eq. 8 $$k^- = k^+ exp\left(-\frac{\Delta W}{k_B T}\right)$$

956   where ΔW is the energy penalty for detaching a growth unit from a kink site to the solution or

957   vapor (the total work of detachment from a kink site). For both Kossel growth units and

958   centrosymmetric molecules ΔW is the same for every site (although the kink, edge and terrace

959   energies are different for different edges[214]). This explains why u is not edge-dependent for such

960   crystal systems, and leads to a linear relationship for u as a function of relative supersaturation ($S-$

961   1)[215]

962   Eq. 9 $$u = k^+ x_{sat}(S - 1)$$

963   and also a linear relationship for step velocity

964   Eq. 10
$$v_i^{step} = a_{p,i}\rho_i k^+ x_{sat}(S-1)$$

965   This linear step velocity relationship has been validated by experiment.[216-218]

966   This mechanistic modeling framework is completed once the step edges are identified on each

967   candidate crystal face. This is done using the concept of Periodic Bond Chains (PBCs).[219-221]

968   These correspond to the directions of the strongly bonded growth units throughout the 3D solid

969   structure. Faces that have two or more PBCs generally grow by either the spiral mechanism or the

970   2D nucleation mechanism where the edges of the spirals or 2D nuclei correspond to the PBC

971   directions.[203] Such faces grow slowly relative to other faces (which have only one or zero PBCs[219-

972   221]) and therefore dominate the surface structure of the crystal, as first stated by Gibbs in 1875[222]

973   (footnote on pp. 325-326).

974   The overall mechanistic approach is depicted schematically in Figure 10.
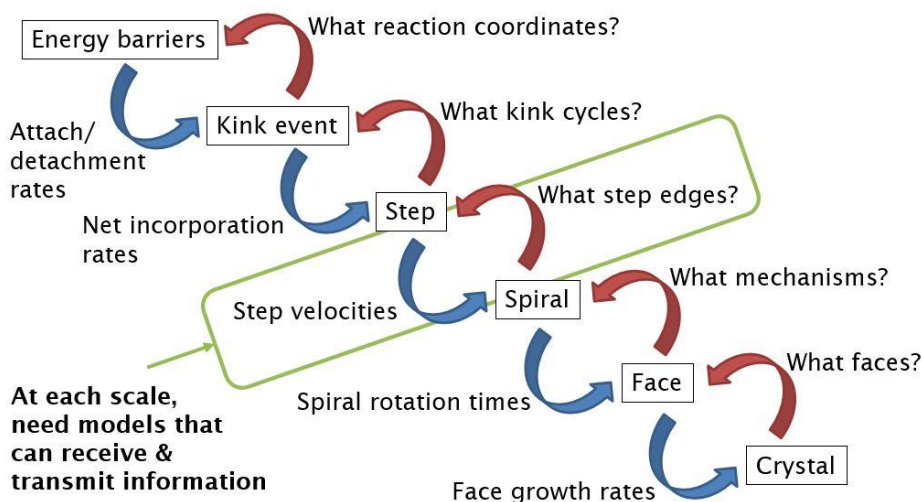
975

976   Figure 10. Schematic representation of the mechanistic growth model for calculating absolute

977   growth rates of crystal faces. The red arrows indicate the questions asked during assembly of the

978   growth model; the blue arrows indicate the type of calculations required to execute the growth

979   model.

980   If absolute growth rates are desired, then this mechanistic modeling approach requires that we

981   calculate values for x_{sat} and k^+. Saturation composition can be calculated by the methods described

982   in Section 4 and the attachment rate constant as discussed in Section 5.3. However, morphology

983   prediction depends only on the relative growth rates of the crystal faces, R_f, (where the index f

984   refers to a specific crystal facet among a set on the crystal surface) where it is conventional to use

985   the slowest growing face (let's say face 1) that dominates the morphology as the reference, $R_f = G_f$

986    $/ G_1$. For centrosymmetric molecules the quantities x_{sat} and k^+ cancel from the relative growth

987   rates and thus do not need to be known. The steady-state growth shape of the crystal is determined

988   from the Frank-Chernov construction[223, 224]

989   Eq. 11
$$\frac{R_1}{x_1} = \frac{R_2}{x_2} = \cdots \frac{R_f}{x_f} = 1$$

990   where x_f now represents the relative perpendicular distance of face f to an origin at the center of

991   the crystal (relative to the absolute perpendicular distance of face 1 to the same origin). A dynamic

992   evolution model can also be computed for faceted crystals starting from an arbitrary initial faceted

993   shape containing all the likely low index crystal faces (such as, e.g., a "disco ball") which

994   ultimately leads to Eq. 11 at steady-state.[225, 226] The overall mechanistic approach for predicting

995   morphology is depicted schematically in Figure 11 and reviewed by Li *et al.*[227]
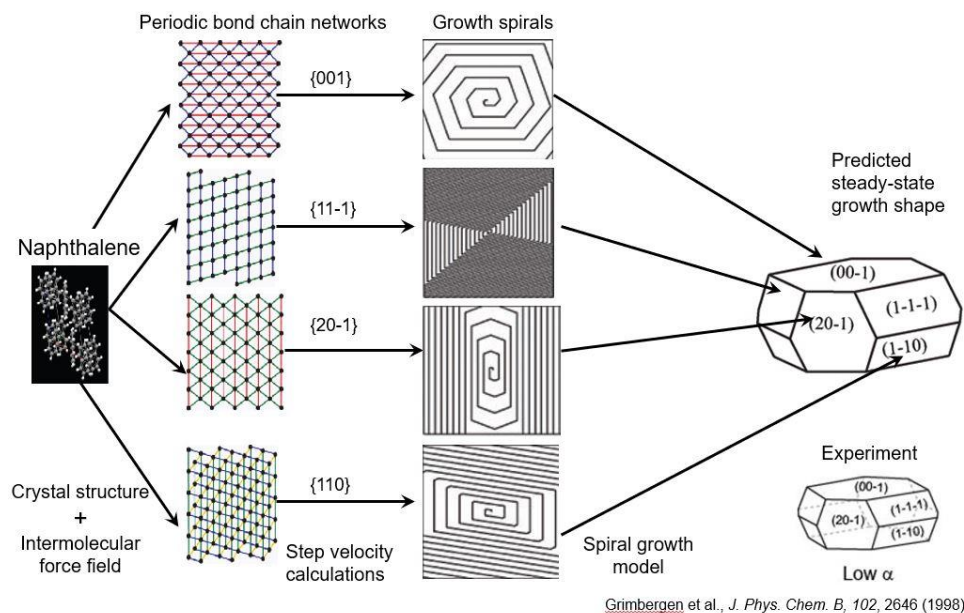
Figure 11. Schematic representation of the mechanistic growth model for predicting crystal morphology, applied to the growth of naphthalene from ethanol solution.[203, 228] Adapted from Figure 5.1 in Li *et al.*,[227] with permission from Elsevier.

It is best to start morphology studies by measuring and predicting the sublimation enthalpy and the morphology for sublimation-grown crystals. This eliminates all solvent effects and allows for:

1. a fidelity test of the chosen classical force field to represent the solid state. If there is a big difference between the predicted and measured sublimation enthalpy, then it is highly likely that the kink energy (and other bond energies) in the solid are in error. In such a case it is unlikely that the morphology predictions will be good. As an example, GAFF predicts a sublimation enthalpy for adipic acid of approx. -170 kJ mol^{-1} whereas the measured value is approx. –130 kJ mol^{-1}. The resulting morphology predictions are poor.

2. comparison of the morphologies calculated with different solvents to the "reference" sublimation shape, and hence a qualitative assessment of the solvent effect, i.e., how

51

1010   different functional groups on the solvents interact with various functional groups exposed

1011   on the different crystal faces.

1012   5.1   Examples for centrosymmetric molecules

1013   Tilbury et al.[212] report just such a study for four centrosymmetric crystal systems, namely,

1014   experimental and predicted morphologies for both sublimation-grown crystals and solution-grown

1015   crystals for biphenyl, adipic acid, pentaerythritol and naphthalene. In each case the solute

1016   molecule is unusually symmetric with an inversion center. The Gavezzotti Coulomb-London-

1017   Pauli (CLP)[229] force field was used for the solid state energy calculations and the van Oss,

1018   Chaudhury and Good (vOCG)[230] interface model to capture the solvent effect. In each case the

1019   predicted morphology is in good agreement with the experimental morphology, and in each case

1020   there is a significant shape change between sublimation-growth and solution-growth. For

1021   example, adipic acid grows as a rod from the vapor but as a hexagonal plate from aqueous solution

1022   (Figure 12).
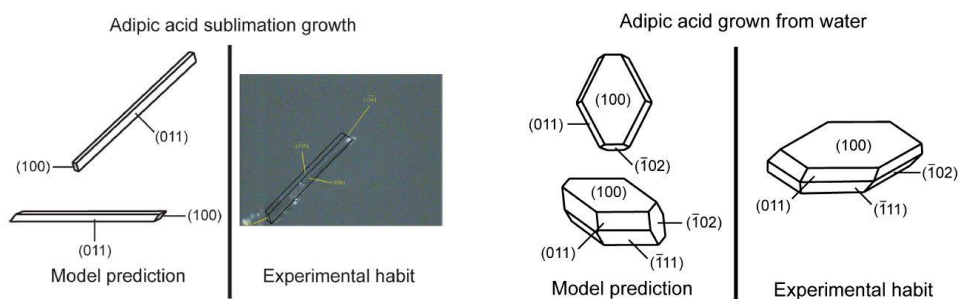
1023

1024 Figure 12. Adipic acid crystals. (a) two views of the predicted morphology for sublimation growth

1025 (left) and the experimentally grown sublimation shape (right), (b) two views of the predicted

1026 morphology for growth from aqueous solution (left) and a rendering of the experimental shape

1027 from aqueous solution (right).[231] Adapted from Figure 6 and Figure 7 in Tilbury et al., [212]

1028　　　In Figure 12, it is highly noticeable that the (100) face is the smallest face on the sublimation

1029　　　shape but the largest on the solution-grown shape. Thus, the presence of water transforms the

1030　　　(100) from the fastest growing to the slowest growing face. This face exposes acid groups, and in

1031　　　particular O-H groups which form hydrogen bonds with water. Thus, the face becomes strongly

1032　　　solvated with water molecules which increases the solvation activation barrier for detachment of

1033　　　water molecules from kink sites (or equivalently, attachment of adipic acid growth units to kink

1034　　　sites) on the (100) surface (see section on absolute growth rates 5.3). The result is a significant

1035　　　reduction of the face growth rate relative to sublimation growth. Davey *et al.*[231] described the

1036　　　process being modeled as "adipic acid molecules approaching the (100) face do not experience the

1037　　　expected H-bonded interaction with molecules in the surface. One reason for this could be that

1038　　　the surface is covered by a layer of water molecules occupying the H-bonding sites such that

1039　　　growth of this surface is controlled by desorption of water molecules rather than attachment of

1040　　　adipic acid molecules."

1041　　　Two other notable examples of morphology prediction for centrosymmetric growth units are

1042　　　shown in Figure 13, for succinic acid grown from aqueous solution, and for olanzapine grown

1043　　　from acetone. Olanzapine molecules are not centrosymmetric, but it has been established that the

1044　　　building block of most known forms of olanzapine (3.1.2) incorporate into the crystal lattice as

1045　　　centrosymmetric molecular dimers – thus the growth unit is centrosymmetric.[78, 232-235]
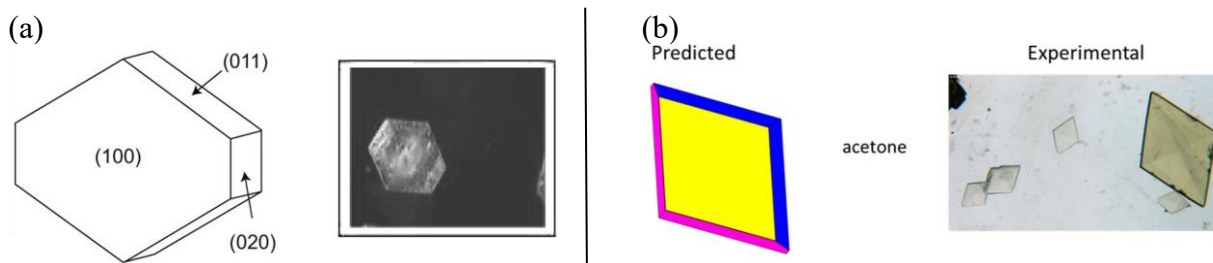
1046

1047   Figure 13. (a) Succinic acid crystals (β form) grown from aqueous solution; left – predicted

1048   morphology, right – experimental shape, adapted from Figure 7 in Snyder *et al.*,[226] with permission

1049   from Wiley, Copyright © 2007 American Institute of Chemical Engineers (AIChE), and (b)

1050   olanzapine crystals (Form I) grown from acetone; left - predicted morphology (yellow=(100),

1051   blue=(11-1), magenta=(011)), right – experimental shape, adapted from Figure 7 in Sun *et al.*[215]

1052   5.2   Growth and morphology for non-centrosymmetric molecules and high supersaturations

1053   Crystal morphology prediction for simple centrosymmetric molecules is well in hand and quite

1054   reliable for a wide selection of crystal systems, especially at low levels of supersaturation.

1055   Extending the methods to high supersaturation (typical of antisolvent crystallization) and to non-

1056   centrosymmetric growth units (e.g., most API molecules and all ionic crystals) remain areas of

1057   current research. Good progress has been made on both these topics, but the theory is incomplete

1058   for crystals under high supersaturation with non-centrosymmetric growth units, e.g., antisolvent

1059   crystallization of API molecules.

1060   It is known from kinetic Monte Carlo (kMC) simulations of Kossel crystals that kink density

1061   increases rapidly as supersaturation increases, thus the Boltzmann treatment for this quantity

1062   becomes invalid.[236, 237] To establish the correct supersaturation dependence of kink density, a

1063   master equation model is constructed that balances the rate of kink-forming events with the rates

1064   of kink-destroying events. In the context of a Kossel crystal model, the "gold standard" work on

1065   this topic was published twenty years ago by Cuppen *et al.*[236] where they derived the non-

1066   equilibrium kink density model for kinks of all heights and showed that it agrees perfectly with

1067   kMC simulations over a wide range of supersaturations, and moreover agrees analytically with the

1068   Boltzmann equilibrium kink density expression (also derived in the paper) for kinks of all heights.

1069   They also showed that the step velocity model resulting from the new kink density model agrees

perfectly with kMC simulations and is nonlinear over a wide range of supersaturations. Thus, crystal growth and morphology models for simple centrosymmetric molecules are well in hand even at high supersaturations, although we do not know of any direct tests with experimental morphologies for crystallizations at high supersaturation.

This leaves growth and morphology modeling for non-centrosymmetric molecules (at both low and high supersaturation) as the key targets of current research. Kink rate models have been developed for two growth units[238] and for any number of growth units.[214, 239] One of the new phenomena that arises when there are multiple growth units in the unit cell is the pattern of growth units in the rows of steps. Take the case of two growth units, A and B, in the unit cell, where A and B may stand for two ions (e.g., Na^+ and Cl^-), or two non-centrosymmetric molecules which are arranged in different orientations in the unit cell (e.g., in a herringbone pattern). In some crystallographic directions the steps have a row of AAAAAA growth units followed by a row of BBBBBB, then repeating; while in other crystallographic directions the pattern is ABABAB along every row (other patterns also occur). There is no such behavior for Kossel or centrosymmetric growth units. Among the consequences of this are:

1. steps in different crystallographic directions on the same face (i.e., different edges of the same spiral or 2D nucleus) grow by different kink rate mechanisms that lead to different step velocity models.

2. steps that have different rows of growth units (e.g., AAAAAA followed by BBBBBB) have different edge energies for the different rows. As a result of anisotropic interactions, one of the rows can become unstable (thermodynamically unfavorable to elongate) and this behavior must be taken into account in the growth modeling.[214, 239-241]

1092   These concepts have been generalized to any number of growth units in the unit cell, leading to

1093   the concept of kink cycles (and corresponding kink rate formulae) that capture the various

1094   patterns.[214, 239] However, the link between step velocity and kink rate/kink cycle for crystals with

1095   multiple growth units is still not settled and multiple relationships have been proposed.[214, 239, 240]

1096   When these methods are put in the hands of experts, good results for predicted crystal morphology

1097   may be obtained for a wide range of complex crystals with non-centrosymmetric growth units.

1098   Especially notable are the papers by Shim and Koo.[242-244] Nevertheless, it remains a challenge to

1099   convert physics-based expert knowledge to a general-purpose digital design aid for morphology

1100   prediction that is robust enough to use in API workflows. The two most advanced digital aids at

1101   the present time are ADDICT[245] and CrystalGrower.[246]

1102   Two successful examples of morphology predictions for crystals with non-centrosymmetric

1103   growth units are shown in Figure 14.

1104
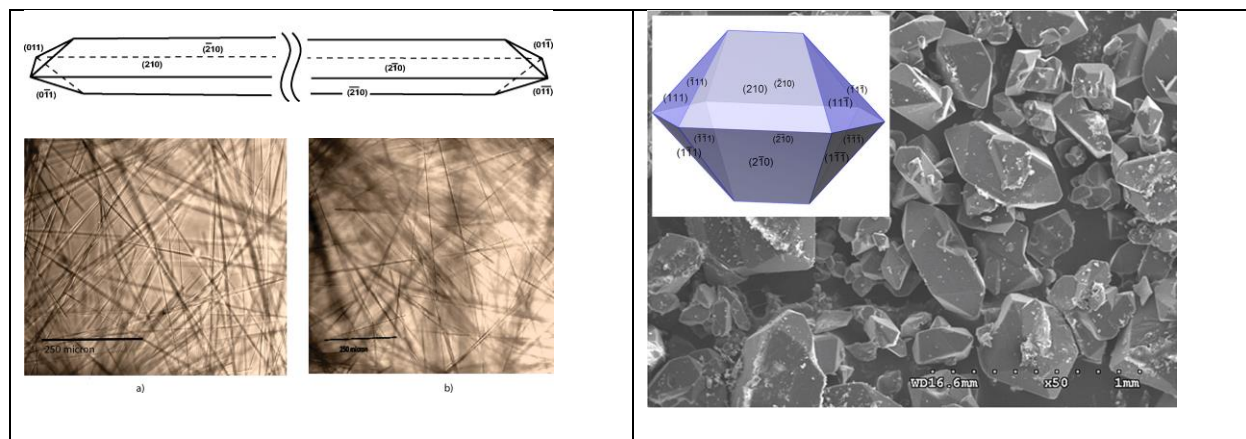
1105 Figure 14. Left: lovastatin crystals grown from alcohol solutions; top – predicted morphology,

1106 bottom – experimental shape grown from isopropanol (left) and methanol (right), adapted from

1107 Figures 26 and 27 in Kuvadia *et al.*[239] Right: the explosive RDX grown from acetone with

1108 predicted morphology inset, adapted from Figures 6 and 8 in Shim *et al.*[242]

1109   Improved methods for estimating solvent effects may lead to better morphology predictions. A

1110   new approach that uses a "thermodynamic path" to obtain an expression for the solvent-modified

1111   bond energy between two neighboring growth units C & D on a crystal step edge is given by the

1112   equation:

1113
$$\Phi_{CD}^{solv.modif.}(liq) = \Phi_{CD}^{vapor}(solid\ state) + \Delta G_{monomer}^{solv} - {\Delta G_{BP}^{solv}}/{2}$$

1114   where $\Phi_{CD}^{solv.modif.}(liq)$ is the solvent-modified bond energy between growth units C & D in

1115   presence of a solvent, $\Phi_{CD}^{vapor}(solid\ state)$ is the bond energy between growth units C & D in a

1116   vapor environment (i.e., without solvent present, and is calculated directly from the atom- atom

1117   force field in the solid state), $\Delta G_{monomer}^{solv}$ is the solvation free energy for the monomer growth unit

1118   when immersed in the solvent, and ΔG_{BP}^{solv} is the solvation free energy for the bonded pair of

1119   growth units C & D when immersed in the solvent. These solvation free energies may be estimated

1120   using molecular simulation or COSMO-based methods.

1121   5.3   Absolute growth rates

1122   We now return to the matter of predicting absolute growth rates of crystal faces. The key

1123   quantity needed is k^+, the attachment rate constant for growth units into kink sites. This is

1124   calculated using molecular simulation to generate the free energy landscape at a given temperature

1125   and pressure for the undocked and docked growth unit in the kink site (i.e., kink site occupied by

1126   solvent and solute, respectively). The transition path through this surface (locus of minimum free

1127   energy) is determined and the activation energy barrier is calculated. Transition state theory is

1128   then used to obtain a value for k^+. A major success using this approach was reported[247] for barite

1129   (barium sulfate) grown from aqueous solution. The calculated free energy profile as a function of

1130   reaction coordinate is shown in Figure 15, where there are four separate transition states (peaks in

1131   the graph) between five energetically distinct states (valleys). Repeating the calculations at various

1132 temperatures allows for the calculation of the rate-limiting attachment and detachment rate

1133 constants as a function of temperature, which yield a straight line on an Arrhenius plot (see Fig. 4

1134 in ref [247]). The resulting activation energies for attachment and detachment are in good agreement

1135 with experimental estimates of the activation energies for step growth and dissolution,

1136 respectively.
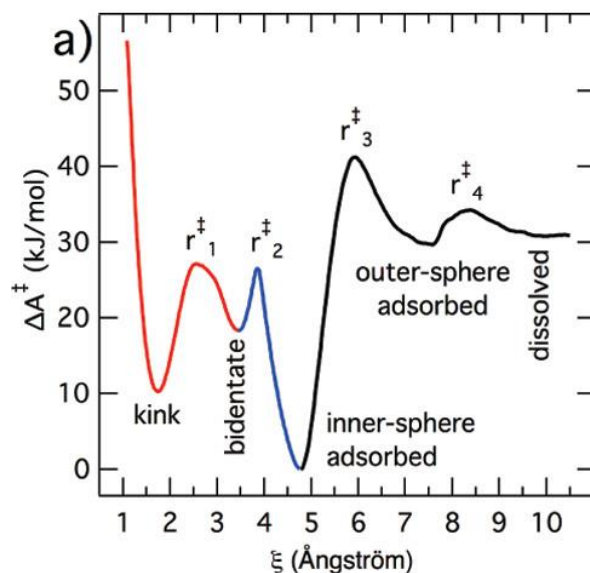
1137

1138 Figure 15. Calculated free energy profile for barium ions as a function of reaction coordinate for

1139 barite grown from aqueous solution, reproduced from Figure 2 in Stack *et al.*[247]

1140 Joswiak *et al.*[166] performed a similar study for the simpler system of sodium chloride

1141 crystallization from aqueous solution (Figure 16). The simplicity of this system permitted a

1142 prediction of absolute crystal growth rate as a function of supersaturation (Figure 16c).[248] All the

1143 major crystal faces of sodium chloride ((001), (010), (001)) are identical, as are the steps on each

1144 face, thus limiting the number of calculations required to perform the growth rate calculation.

1145 Once a value for k^+ was calculated it was inserted into their equation for calculating the crystal

1146 face growth rates to obtain Figure 16c.[248] The resulting predictions are in good agreement with

1147 experiment.
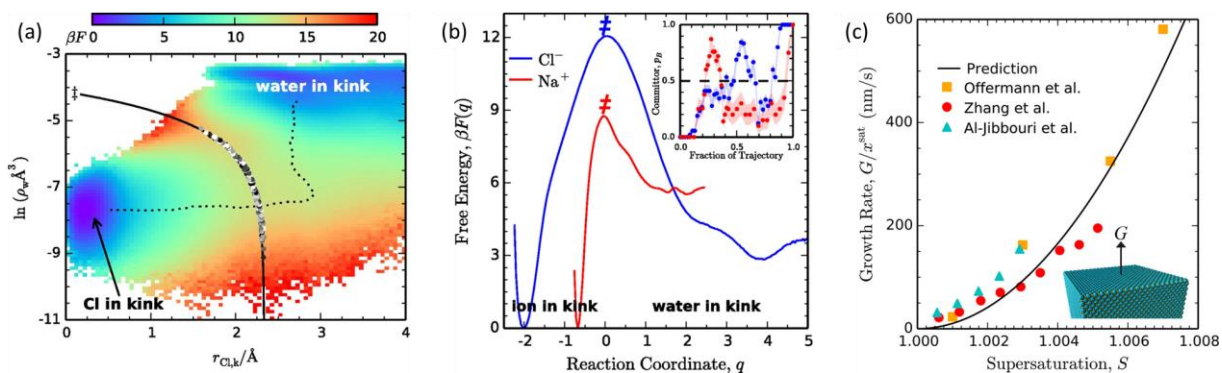
Figure 16. (a) Free energy landscape for chloride kink site on a sodium chloride step. The CVs are the distance of the chloride ion from the kink site and the local water density in the kink site. The minimum free energy path (dotted) indicates that the chloride attachment process is two-step; first water molecules leave the kink site (the site desolvates) then the chloride ion docks in the cavity created. (b) Free energy profile as a function of reaction coordinate. Reproduced from Figures 3 and 4 in Joswiak *et al.*[166] (c) Comparison between predicted NaCl crystal growth rate (solid black curve) and experimental measurements (colored symbols), reproduced from Figure 4 in Joswiak *et al.*[248]

Recent molecular simulation studies on solvation effects and activation barriers in crystal growth from solution include glutamic acid crystallization from aqueous solution, and concluded that solvent fluctuations in the solvation shell are among the factors determining the activation barrier for crystal growth rates.[249]

One useful byproduct is that these state-of-the-art modeling techniques enable the prediction of absolute face growth rates from the predicted relative growth rates plus one absolute growth rate (for the reference face) from either experiment or molecular simulation. This enables a big reduction in experimental effort to obtain such growth rates. The resulting growth rates can then be used, for example, in multidimensional population balance models, as noted by Kuvadia and Doherty.[250]

1167   5.4   Open problems for morphology and growth rate

1168   Despite these successes, modeling methods are not quite ready for application to the everyday

1169   workflow in API research and development (but they are much closer than they were ten years

1170   ago). Open problems include:

1171   • developing non-equilibrium kink density models (at high supersaturation) for crystals with

1172   many (> 2) non-centrosymmetric growth units in the unit cell.

1173   • developing step velocity models for such systems.

1174   • improving methods for estimating solvent-modified bond energies, particularly for solvent

1175   mixtures, and especially mixtures of solvent and antisolvent.

1176   • prediction methods for attachment rate constants for API-like solutes with multiple

1177   "floppy" functional groups. It has been shown that conformational equilibria as well as

1178   conformational transition mechanisms can become the kinetic bottleneck in crystal

1179   growth,[251] and that the presence of slow conformational equilibria can affect overall

1180   crystal growth rates at process scale.[190] How to identify molecular systems that are subject

1181   to these conformational limitations and how to generalize growth rate estimates in such

1182   cases remains an open challenge.

1183   6   Nucleation

1184   Nucleation is the initial irreversible step in the formation of a new crystal, so it has an unremitting

1185   influence on later stages of crystallization. Nucleation is a notoriously rare event. The fastest

1186   measurable rates of solute precipitate nucleation, ca. 10^{10} cm^{-3} s^{-1}, are still excruciatingly slow

1187   compared to accessible simulation time scales. To see even one nucleation event at realistic

1188   conditions in a typical simulation box of size ca. 100 nm^3, one would need to simulate a staggering

1189 10^9 s of dynamics. Real simulations are (at present) limited to the microsecond time scale, so

1190 direct simulations cannot access realistic nucleation rates. But a systematic use of direct

1191 simulations can reveal important trends in the nucleation rate as a function of solute

1192 supersaturation,[161, 252, 253] solvent composition,[254-256] surfactants/additives,[256-259] and

1193 characteristics of heterogeneous nucleation sites and surfaces.[260]

1194 The key reason for concern about nucleation is in polymorph discovery. Late appearing

1195 polymorphs are clearly those whose nucleation has usually been out-competed by other

1196 polymorphs. The cases of disappearing polymorphs are likely to arise when the more stable form

1197 has a very slow nucleation rate but fast growth rate (Table 1 in Section 8.3). In this context,

1198 atomistic simulations, while hardly able to provide estimates of absolute nucleation rates, could

1199 help in ranking the relative nucleation kinetics of different forms, at least in the absence of

1200 impurities, and thus provide approaches beyond thermodynamics to refine CSP-generated crystal

1201 energy landscapes and identify long-lived putative polymorphs. An ability to model nucleation

1202 could allow the assessment that there is little risk of a computationally more stable crystal structure,

1203 that has not been observed in extensive screening, disrupting the manufacture of a metastable but

1204 kinetically favored form.

1205 Rare events methods[101, 261, 262] made it possible to compute single-component nucleation rates at

1206 realistic conditions, even for extremely slow nucleation processes.[263-266] However, formidable

1207 challenges remain, especially for solute precipitate nucleation[267] and heterogeneous nucleation.[268]

1208 The challenges originate from multiple factors.

1209 • Unlike predictions of structure and stability which only require bulk solid properties,

1210 understanding nucleation requires a quantitative understanding of the solvent properties,[254,

1211 256, 269] interfaces,[270, 271] and perhaps also interactions with other surfaces, impurities,

1212 surfactants, etc.[272, 273]

- Nucleation rates are highly sensitive to supersaturation. Except in highly coincidental cases where the force field exactly predicts the solubility limit, a simulation is not being performed at the experimental supersaturation concentration. For quantitative comparisons and predictions, experimental and simulated supersaturations should match. Therefore, precise solubility and driving force calculations are important (and non-trivial) first steps.[154, 274, 275]

- Nucleation is an intrinsically irreversible non-equilibrium process, and (for solute precipitate nucleation) the solid invariably has a different composition from the starting solution.[267, 276] The transformation from a solution with one composition to a solid with another poses tremendous difficulties for simulations.

- Most theoretical analyses and simulations focus on well-defined but improbable homogeneous primary nucleation processes.[277] However, most nucleation processes occur by heterogeneous nucleation[269, 272, 278, 279] or secondary nucleation.[280, 281] This daunting challenge motivates much of the discussion below.

Theory and molecular simulations have overcome numerous hurdles, but accurate predictions remain elusive for all but the simplest of systems. The following sections highlight key developments in primary homogeneous nucleation, primary heterogeneous nucleation, and secondary nucleation models as related to pharmaceutical crystallization and polymorph screening.

6.1   Homogeneous nucleation

Homogeneous nucleation, in the context of API crystallization, is the spontaneous emergence of a stable crystallite or amorphous precursor from the pure metastable solution, with no assistance from templating surfaces or particulates. In experiments, homogeneous nucleation tends to be the most difficult pathway to observe, because heterogenous (Section 6.3) or secondary nucleation

1236   (Section 6.2) pathways usually intercede.   For simulations, understanding homogeneous

1237   nucleation is often the first goal because it occurs in a well-defined pristine solution. Nevertheless,

1238   even homogeneous nucleation poses serious difficulties. In particular, nucleation studies require

1239   efficient advanced sampling schemes to overcome the rare events problem,[161] precise solubility

1240   calculations to match simulated and experimental supersaturations,[154, 162, 274, 282-284] accurate order

1241   parameters for cluster size calculations,[254, 255] surface free energy-corrected nucleation rate

1242   estimates,[285, 286] and methods to predict the impact of surfactant adsorption.[257, 258, 287] There is a

1243   significant problem in controlling the supersaturation during a simulation of solute precipitate

1244   nucleation, because of the difficulty of keeping the chemical potential difference constant in a

1245   small box with constant number of molecules.[288] Putting a small nucleus to seed the simulation

1246   can estimate rates without artifacts from supersaturation depletion, but this requires *a priori*

1247   assumptions about nucleus structure.[274, 275, 289, 290]   Osmotic ensemble methods can control

1248   supersaturation with no assumptions about nucleus structure,[291] which may prove particularly

1249   useful in studies of two-step homogeneous nucleation.[271, 292, 293]

1250   6.2   Secondary nucleation

1251   Secondary nucleation, in multiple ways, is the opposite of homogeneous nucleation. While

1252   homogeneous nucleation involves molecular-scale spontaneous assembly from solution with no

1253   crystals, secondary nucleation occurs by mechanical breakage of existing macroscopic crystals

1254   into pre-formed viable crystallites.[267, 294, 295] Homogeneous nucleation rarely (if ever) occurs in an

1255   industrial crystallizer, while secondary nucleation is thought to be the dominant mechanism.[280, 296-

1256   298] Quantitative models have been developed for secondary nucleation, assuming that crystals can

1257   fracture on collision, accounting for crystal-crystal collisions, impeller-crystal collisions, and other

1258   mechanisms.[267, 294, 295, 299] The solute morphology and mechanical properties[300] are important in

1259   determining the ease of fracture, but this is not explicit in current models, in which fracture enters

1260   through an experimental parameter. Secondary nucleation models are extremely useful for

1261   industrial crystallization,[301] but not for polymorph screening, where the main targets are low

1262   energy solid forms that are predicted be kinetically stable, but are yet-to-be experimentally realized

1263   (Section 3.3).

1264   6.3   Heterogeneous nucleation

1265   Heterogeneous nucleation can occur where the solution contacts a bubble, an oil droplet, or an

1266   atomically smooth solid surface.[267, 273, 302] Like homogeneous nucleation, these special cases of

1267   heterogeneous nucleation are amenable to theoretical modeling and molecular simulations. We

1268   have gained tremendous insight from these studies, e.g. how solid-crystal, solid-solution, and

1269   crystal-solution interfaces modulate nucleation barriers[272, 273, 279] and how barriers for

1270   heterogeneous nucleation relate to barriers for homogeneous nucleation.[279, 303] Theory and

1271   simulation have also explained how barriers are influenced by additional factors like line

1272   tension,[304, 305] lattice mismatch/elasticity,[306, 307] curvature,[308, 309] hydrophobicity,[310, 311]

1273   electrostatics,[312, 313] and adsorbates.[309, 314-316]

1274   6.3.1   Heterogeneous nucleation for polymorph discovery

1275   In a few impressive cases, instead of using CSP output to select an isomorphous crystal (as in

1276   Section 3.3), special epitaxially matching surfaces have been identified and prepared to promote

1277   heterogeneous nucleation of a desired form.[151, 152, 317-320] Since this is a surface phenomenon, the

1278   use of natural ledges[321] or even engineered surfaces[322] can be worth trying.[323] Both the shape and

1279   the nature of the exposed functional groups appear to be important in designing surfaces for

1280   polymorph discovery or control.[324, 325] However, there are many situations, such as the discovery

1281   of the γ form of succinic acid in a failed purification by co-crystallization experiment,[82] or the

1282   stabilization of a form of progesterone by impurities,[326] where it has been impossible to identify

1283   the specific impurities or templates involved. Because identifying and preparing ordered structure-

1284   selective templates is not trivial,[319] template discovery and fabrication will probably be pursued

1285   only for otherwise elusive targets of the highest priority.

1286   6.3.2   Heterogeneous nucleation on disordered nucleants

1287   Heterogeneous nucleation is not limited to perfect surfaces. It can also occur in pores, pits, and

1288   other defect sites.[327] Theory and simulation have also provided important and generalizable

1289   insights about heterogeneous nucleation in these environments, e.g. the effects of pore size[328] and

1290   wedge angles[325] on nucleation rates. Again, the pore environments and geometries in these studies

1291   are somewhat idealized. The story of heterogeneous nucleation will be far more complicated for

1292   real scratches, cracks, fissures, and defects.[327, 329]

1293   In addition, experiments[49, 265, 330] and calculations[48, 331, 332] show that nanopore confinement can

1294   reverse the relative stabilities of polymorphs and hydrates because the small size of the crystallites

1295   makes the surface energy important. Because different pore sizes and geometries can favor

1296   different polymorphs, porous disordered media are extremely useful for polymorph screening.

1297   "Naomi's nucleants"[333] are porous bioglass materials that promote nucleation of protein crystals.

1298   Polymers are also used, with porous polymer films being successful in numerous polymorph

1299   selection efforts,[334-338] and polymer dispersions producing olanzapine form IV.[79]

1300   To understand the advantages of disordered media for promoting nucleation, one must remember

1301   that nucleation is irreversible once the particle is large enough. In a small sample, the first event

1302   spurs the formation of a growing crystal that then consumes the supersaturation and suppresses

1303   subsequent nucleation events. Accordingly, the induction time and the resulting polymorph are

1304   the result of extreme value statistics. The most potent nucleation sites among thousands in the

65

1305 disordered media will determine the outcome. This is true for each droplet in a nucleation assay,

1306 but the most potent nucleants differ from droplet to droplet[339] which make these assays useful in

1307 solid form screening. Once the first crystal of a new polymorph has been found, then it can be used

1308 to seed more conventional crystallization experiments, provided the growth and transformation

1309 kinetics are favorable.

1310 6.4 Open problems in nucleation

1311 There is some debate about the extent to which nucleation is a problem for pharmaceutical

1312 manufacturing. In most industrial crystallizers, nucleation can be practically avoided by seeding,

1313 and the collision-induced mechanism of secondary nucleation mainly eliminates the need for

1314 molecular level understanding of primary nucleation processes. However, at the polymorph

1315 screening and solid form discovery phase, nucleation is critical. In academic studies aimed at

1316 fundamental molecular scale understanding of nucleation, the diversity of heterogeneous

1317 nucleation sites and mechanisms is a frustrating nuisance. In contrast, polymorph screening and

1318 solid form discovery efforts could potentially exploit the number of sites that can promote

1319 heterogeneous nucleation[340] and their promotion of different forms to different degrees. Thus,

1320 there is a need for empirical statistical analyses that can help design better polymorph screening

1321 assays, along with research aimed at understanding molecular level nucleation mechanisms.

1322 A key issue will be the extent to which the simulation methods and analysis tools for simple

1323 systems translate to API molecules. The dominant mechanisms for conformational transformation

1324 can differ significantly between the solution phase and as the molecule becomes incorporated into

1325 the crystal, even for molecules as small as ibuprofen.[251] An early systematic study aimed at

1326 quantifying the impact of molecular flexibility on the nucleation kinetics of four para substituted

1327 benzoic acids suggests that conformational change is only one of the many activated processes that

1328   contribute to and control the kinetics of crystallization. Its relative weight in the whole process

1329   will change depending on the system.[341]

1330   ## 7   Process Design

1331   The digital design of crystallization processes requires translating the insight obtained by

1332   studying crystallization at the atomistic scale into process-level information. This is achieved by

1333   developing population and material balance models of crystallizers. Morphology, growth rates,

1334   solubility and often nucleation rates are used in population balance models which predict the

1335   evolution of the particle size distribution (PSD) and, in the case of multidimensional population

1336   balance models, the evolution of the crystal shape distributions of the crystalline product. These

1337   outputs are essential for the design of crystallization processes. For example, the predicted PSD

1338   may be compared with the desired target PSD to determine whether wet milling, dry milling, or

1339   wet granulation operations should be included in the process. We construct a mixed suspension

1340   mixed product removal (MSMPR) crystallizer model, applicable for batch, semi- or fed-batch and

1341   continuous processes. This model assumes that the crystallizer provides perfect mixing with no

1342   gradients in solute concentration, crystalline suspension, temperature or anti-solvent. The

1343   population balance for an MSMPR assumes the crystallizer has an inlet stream(s) that supplies a

1344   fresh solution, a stirred tank containing a suspension of crystals in solution, and an exit stream that

1345   carries the suspension out of the vessel. In its most general form, the population balance model for

1346   such a system is given by:

1347   Eq. 12
$$\frac{\partial(nV)}{\partial t} + \frac{\partial}{\partial L}(GnV) = V(B - D) - qn + q_{in}n_{in}$$

1348   where $n(\mathbf{L}, t)$, is the population density of particles per unit volume of suspension, $\mathbf{L}$ is a vector of

1349   characteristic crystal dimensions in the different crystallographic directions derived from the

1350   crystallographic (hkl) values of the morphologically-relevant faces, t is time, V is the suspension

1351 volume, q is the volumetric flowrate of the inlet (subscript in) or outlet (no subscript) streams, $\mathbf{G}$

1352 $= d\mathbf{L}/dt$ is the vector of absolute growth rates for each face (from Section 5), and B and D are the

1353 birth and death rates of crystals per unit volume of suspension, e.g. by breakage, aggregation or

1354 agglomeration. In this formulation, the crystallizer volume and the inlet and outlet volumetric

1355 flow rates are allowed to vary with time to accommodate such processes as fed batch, co-addition,

1356 antisolvent or other time variant crystallization processes. Nucleation kinetics, including

1357 secondary nucleation, are expressed via a nucleation rate $\mathbf{J}$ which enters the population balance

1358 model via the boundary condition at $\boldsymbol{L} = 0$ discussed below (see Eq. 16).

1359 The solute concentration, C (mass/volume), is determined from a mass balance:

1360 Eq. 13
$$\frac{d}{dt}\left(C(V - V_S) - V_S \varrho_s\right) = q_{in}C_{in} - q\left(1 - \frac{V_S}{V}\right)C - q\frac{V_S}{V}\varrho_s$$

1361 where ϱ_s is the density of the solid, V_S is the volume of solids in the crystallizer, and V_S/V (or ε)

1362 is the solids volume fraction in the crystallizer:

1363 Eq. 14
$$\frac{V_S}{V} = \int d\boldsymbol{L}\, n(\boldsymbol{L}, t)v_{crystal}(\boldsymbol{L})$$

1364 where $v_{crystal}(\boldsymbol{L})$ is the volume of a single crystal with dimensions $\mathbf{L}$.

1365 Often the growth rate used in the population balance (Eq. 12) is modeled with a power law in

1366 the degree of saturation[342-344] or the relative supersaturation:[345-347]

1367 Eq. 15
$$G_j = k_g^j exp\left(\frac{-E_{A,g}^j}{RT}\right)(S - 1)^{g_j}$$

1368 where the relative supersaturation is defined as $S = C/C_{sat}$, and j refers to the specific crystal face.

1369 To fully characterize the operation of an MSMPR crystallizer the system composed of the

1370 coupled material and population balances (Eq. 12 to Eq. 15) need to be solved, together with initial

1371 conditions and boundary conditions. The initial conditions are:

1372 Eq. 16
$$n(\boldsymbol{L}, t = 0) = n_{seed}$$

1373 Eq. 17 $$C(t = 0) = C_0$$

1374 where n_{seed} is the PSD of seed crystals inside the crystallizer at t=0, and C_0 is the initial

1375 concentration of clear liquid inside the crystallizer at t = 0.

1376   A boundary condition is also required to account for nucleation of particles of microscopic size.

1377 For the mono-dimensional population balance case, when particles can be characterized by a single

1378 length, and vector **L** reduces to a scalar, L, the nucleation boundary condition at L=0$^+$ is:[348]

1379 Eq. 18 $$n(0^+, t) = \frac{J}{G}$$

1380 where 0$^+$ indicates that nuclei have a finite albeit extremely small size. This size is irrelevant to

1381 the process-design models, and instead of defining these microscopic sizes, one normally simply

1382 approximates them to zero, an approximation that has no significant effect on the shape of the

1383 distribution near the origin.

1384   A proper boundary condition in the multidimensional case requires delta functions or

1385 regularization approximations. It is worth noting here that this boundary condition holds

1386 irrespective of the dominant nucleation mechanism, and either primary or secondary nucleation

1387 can be introduced into crystallizer models following this approach. However, for primary

1388 nucleation, the boundary condition must be modified to account for the fact that **J** represents the

1389 number of nuclei born per unit volume of clear liquid per unit time, rather than per unit volume of

1390 suspension per unit time.[348] Most continuous crystallization processes rely on secondary

1391 nucleation to generate new particles, e.g., the birth of new particles by attrition from existing

1392 crystals due to collisions with walls, impellers, or other crystals.[34, 349-352] For these cases, the *ab*

1393 *initio* conceptional design of a continuous crystallization process or a seeded batch crystallization

1394 process is not possible since secondary nucleation models cannot be predicted *ab initio*. There are

1395 reliable empirical models for secondary nucleation with just a few material- and crystallizer-

1396 specific parameters. The secondary nucleation rate also depends on the stirring rate and/or power

1397 input, the supersaturation, and the solid volume fraction in the crystallizer. For example, a

1398 commonly used model for the secondary nucleation rate[301] is

1399 Eq. 19 $$J = k_N \Omega^2 (V_S/V)(S-1)^i$$

1400 where Ω is the stirrer speed. i and k_N are empirical reactor- and substance-dependent parameters

1401 that must be determined through prior experimentation. These two parameters cannot yet be

1402 predicted from first principles, but identifying the factors that influence secondary nucleation rates

1403 may open doors to fruitful data driven models in the future.[280, 295, 353]

1404   Some continuous crystallization processes use continuous seeding or seed generation,[354] where

1405 the size of the seeds fed to the cascade of crystallizers defines the final shape and crystal size

1406 distribution. These processes can also be modeled using the population balance framework.

1407   Eq. 12 can also be simplified to obtain a model for seeded batch crystallization processes. For

1408 a seeded batch process that operates under conditions where primary and secondary nucleation are

1409 avoided, particle size and shape may be predicted by:

1410 Eq. 20 $$\frac{\partial(nV)}{\partial t} + \frac{\partial}{\partial L}(GnV) = 0$$

1411 together with the solute balance and the initial conditions in Eq. 16 and Eq. 17.

1412   The seeded, batch crystallizer case is relevant for the majority of drug substance crystallization

1413 processes, and for this class, it is possible to consider the *ab initio* conceptional design of the batch

1414 process, provided that solubility can be predicted. This problem was first considered by Mullin

1415 and Nyvlt[355] and later by Ward et al.[356, 357]

1416 7.1   Open Problems in Conceptual Process Design

1417   A major bottleneck in *ab initio* crystallizer design is the current inability to predict secondary

1418 nucleation from first principles. Until this problem is solved there can be no further progress in *ab*

initio crystallizer design. Related difficulties include the need for first principles models for agglomeration and attrition.

While there has been some success at predicting homogeneous nucleation rates and critical nucleus size using a modified classical nucleation approach (employing the Tolman correction to account for the size dependence of surface energy at the nanometer size scale) this has no impact on crystallizer design since homogeneous nucleation is normally irrelevant in industrial practice. This is by design, not by accident. Induction times for homogeneous nucleation are usually so long that they are avoided by seeding crystallizers, whereupon secondary nucleation is the dominant mechanism for producing nuclei. Seeding is desired for control of PSD and crystal form.

Multidimensional population balances have the potential to predict not only the particle size distribution but also the morphology/shape (habit) distribution. However, they are not used widely because a separate growth rate model is needed for each individual facet on the crystal surface (including facets that may ultimately grow off the crystal surface at long time and thus never appear at steady-state). Measured growth rates of individual facets of the same crystal are available for only a handful of systems, and this is unlikely to change in the foreseeable future due to the difficulty of making such measurements. As noted earlier in section 5, one way to overcome these difficulties is to reformulate the multidimensional PBE in terms of a single absolute growth rate for one of the crystal faces along with relative growth rates for all of the others [ref 249 in current version, Kuvadia and Doherty]. *Ab initio* growth models may then be used to estimate all the required growth terms in the PBE. Such an approach has never been tested so we can only speculate about its fidelity and utility in the process design workflow. The approach is ripe for further research.

71

As the conceptual design gets closer to industrial deployment the influence of impurities in the solution cannot be ignored. These are usually present as a result of unwanted byproducts formed during the reaction steps that lead to the API. Some impurities are not recognized by the crystal surfaces (i.e., they do not adsorb) and are inert to the growth mechanism(s). Impurities that do adsorb on crystal surfaces are capable of having a significant effect on growth even at the ppm level (the classic paper by Botsaris et al.[358] provides impressive experimental data for the drastic effect of lead chloride on the growth rate of potassium chloride crystals grown from aqueous solution). Impurities normally reduce growth rates of selected planes thereby modifying the crystal morphology in addition to slowing down the growth process, although instances are known of impurities that accelerate growth.[359-361] There is now a large literature on the effect of impurities and additives on crystal growth; selected papers include those reporting experimental methods and data[218, 231, 362-364] and molecular models.[32, 365-368]

A related issue is impurity sequestration in the crystals, which is also undesired. For pharmaceuticals, where the impurities may be structurally related to the API, but with very different biological effects, it is imperative to avoid their incorporation.[369-371] The usual approach to solving problems caused by impurities is to stage the crystallization by first crystallizing from the crude solution followed by recrystallization from a pure solvent. However, this is not practical for every step along the process route (i.e., for each solid-form intermediate) and there is a push to better understand the effects of impurities with a view to having better control in a single operation.

Higher fidelity crystallizer models that take into account imperfect mixing and other transport effects *via* the use of computational fluid dynamics are available.[372-375] However, they are subject to the same limitations mentioned above with respect to secondary nucleation and impurity effects.

8   Discussion

We have shown how the components of a digital design workflow, going from the molecular structure through to the process design (Figure 2), are in an active state of development and producing worthwhile results. The time will come when this workflow can be integrated into drug discovery, allowing the discovery team to consider not only the medicinal chemistry in selecting candidates to carry forward into *in vivo* studies, but also product and process design and expected bioavailability of the candidate. **This integration will be a game changer.**

This study has illustrated that it is not fully possible for quantitatively reliable simulations to be performed prior to the synthesis of the molecule. The main barrier is the inability to use the same force field to predict the crystal structures and subsequently all the relevant properties and get good enough results. Useful results can be obtained for some properties. Figure 17 shows that the predicted morphologies of succinic acid and olanzapine show little qualitative difference whether the CSP-generated structure or an experimental crystal structure is used, thus demonstrating that morphology can be predicted from the chemical structure. Zhang *et al* demonstrated similar results for ibuprofen grown from several solvents.[376] However, the absolute growth rates do differ, and morphologies can be very sensitive to force field. Similarly, useful guidance on relative stabilities and solubilities for different polymorphs can be achieved, whereas the absolute values are not reliable. This is still extremely valuable information as knowledge of the potential form landscape provides an understanding of the complexity expected in the development of the solid state and oral dosage form. Likewise, a rank order of solubility in different solvents is also extremely useful in guiding solubility screens and solvent selection. From a product development perspective, knowing the expected morphology of a new API provides insight in the complexity expected for drug product development and can help to guide product development towards a continuous direct

1486 compression process for equant morphologies, or a granulation process for needlelike particles.

1487 This is very useful in resource allocation and early phase process planning.

1488 Solubilities, morphologies, and growth rates can be predicted at different temperatures and for

1489 different solvents once the desired form has been produced and some experimental data is available

1490 for that polymorph. The most important limitations on the absolute and relative values of the

1491 properties from computation that have been illustrated by this review are discussed below.

1492
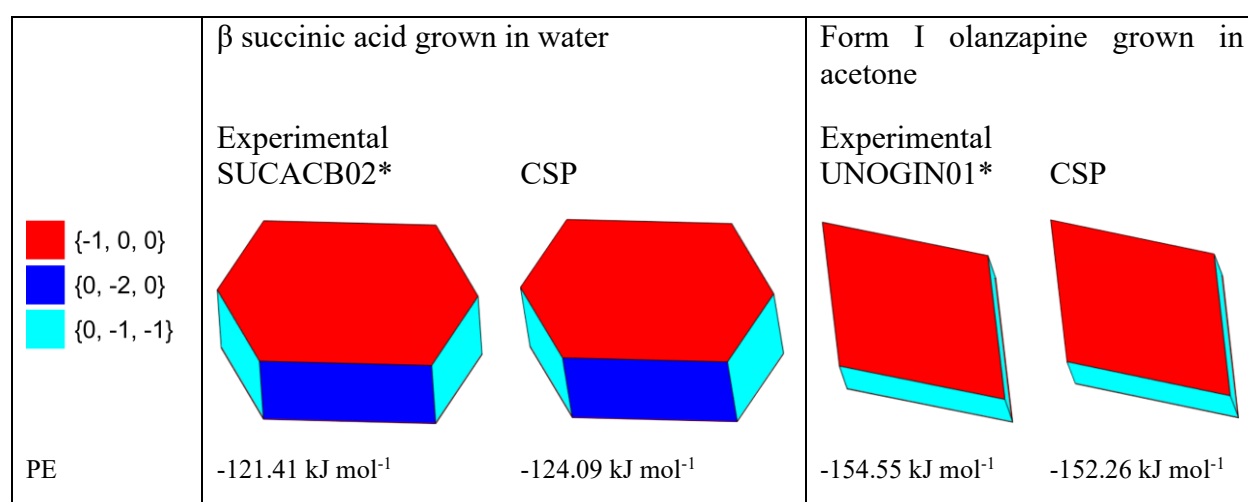
1493 Figure 17. Comparison of morphology predictions using the CSP generated structure and an

1494 experimental structure. The morphologies were calculated with the CLP force field and vOCG

1495 solvent model, and can be compared with other calculations and the experimental morphologies

1496 in Figure 13. The packing energies (PE) calculated with the CLP force field approximate the

1497 lattice energy if conformational changes are neglected and can be compared with other estimates

1498 of the lattice energy for olanzapine in Figure 4 and Figure 5 and with the heat of sublimation of β-

1499 succinic acid in Figure 8 if temperature effects are also neglected. *The experimental structures

1500 had the positions of the hydrogen atoms corrected for the systematic error in X-ray determinations.

1501 Neither structure was optimized with the force field used for the morphologies.

1502 ## 8.1 General problems identified

1503 ### 8.1.1 Crystal structure and temperature effects

1504   In addition to the limitations of the ability of CSP to predict the observed polymorphs, there are

1505 limitations on how well a static perfect crystal structure produced by lattice energy minimization

1506 represents the ambient crystal structure. The effects of thermal expansion are neglected: this can

1507 be very anisotropic, depending on the polymorph and the extent to which there are different

1508 intermolecular forces defining the different cell dimensions. A review of over four thousand

1509 organic crystals shows that the anisotropy of the thermal expansion has a very broad distribution,

1510 with crystallographic evidence suggesting that a third may have at least one orthogonal axis with

1511 negative thermal expansion.[377] Experimentally, the anisotropy shows in the changing appearance

1512 of the pXRD pattern with temperature, with the different peak shifts often changing the peak

1513 overlaps. Computationally, if comparing polymorphs where one has hydrogen bonding in all three

1514 dimensions, and another has hydrogen bonded layers only held together by dispersion interactions

1515 between the layers, the quality of the computational modeling of the structures may be very

1516 different.

1517   In specific cases, the inability of a crystallographic information file to capture the dynamic

1518 structure may limit the accuracy of the calculated thermodynamic and kinetic properties. Dynamic

1519 disorder of even part of the molecule in one polymorph can have a significant effect, but MD

1520 simulations will at least allow diagnosis of whether such effects are likely to be present. Other

1521 limitations include the implicit neglect of defects and unit cell disorder on the properties.

1522   However, it is the curse of the exponential dependence of many properties on temperature that

1523 is the most common limitation in the accuracy of the predicted values. The exponential dependence

1524 is a function of the ratio of the energy differences to temperature, and hence makes computer

1525 modeling very sensitive to the underlying potential energy surface (force field).

1526 8.1.2 Force field availability and accuracy

1527 A major limitation of the digital design of pharmaceutical products is the availability of

1528 sufficiently accurate force fields for pharmaceutical molecules to enable the CSP, landscape

1529 reduction, free energies, morphologies, and other properties to all be calculated to the accuracy

1530 required.

1531 The sensitivity of the pharmaceutical solid state to the underlying potential energy surfaces has

1532 given considerable impetus to academia to develop more accurate energy models for

1533 pharmaceutical crystals.[378] There is a huge effort in the theory, computer codes and testing for both

1534 atomistic force fields[379] (e.g. non-empirical anisotropic atom-atom intermolecular potentials[61]

1535 combined with separated intramolecular force fields[380]), electronic structure calculations[65, 381] and

1536 fragment-based methods.[67] Many of these emerging methods were benchmarked in the 7th Blind

1537 Test of Crystal Structure Prediction.[382, 383] However, there are many issues, ranging from choice

1538 of functional form to selection of experimental or theoretical data for validation, that means that

1539 choice of force field will be a major factor in determining the accuracy of the simulations.

1540 We have noted that the limitations of force field accuracy can to some degree be circumvented

1541 by approaches which use certain properties (e.g. the supersaturation driving force) calculated from

1542 the force field as input to simulations of more complex properties (e.g. absolute and relative growth

1543 rates in Section 5.3).

1544

1545 8.1.3 Computational resources

Many methods described here have only been applied to small molecules, and they have different scaling of computational cost with the size of molecule, its flexibility and the largest unit cell that needs to be considered. The traditional force fields that are used for MD simulations, morphology calculations and some free energy calculations have a functional form that is generally too restrictive for them to be capable of accurately describing pharmaceutical molecules adequately for CSP. Thus, the computational resources required for progress in this field include the optimization of codes that can use more advanced models for the potential energy surfaces. The realism of the implicit assumptions in the traditional force fields, such as using the same parameters for intermolecular and intramolecular terms, assuming isotropic atoms including atomic charges, and ignoring the change in charge density with conformation, will depend on the specific molecules and crystal structures being studied. Commonly available simulation packages can use the same force field for CSP and morphology prediction and give "good enough" results for some molecules for some purposes.[61] Hence, the need for more advanced codes and force fields is molecule dependent. There are hopes that machine-learning of potential energy surfaces may be a route forward,[384] once there is an electronic structure method that can provide an accurate enough potential energy surface for training (c.f. Figure 4). However, the machine-learned force fields also cannot yet be directly used in the highly optimized codes that make long-timescale, large system MD simulations feasible. Thus, a simulation method using a traditional force field that gives high accuracy for benzene may not work well for a flexible API, even if infinite CPU hours could be used.

8.2   Polymorphism and multicomponent systems

The challenges of digital design of crystallization processes are very specific to the API concerned. The kinetic competition between polymorphs is very dependent on the molecule's

1569   structure. What a medicinal chemist would consider minor differences in the molecular structure,

1570   such as introducing a methyl group or changing a hydrogen to a fluorine atom, will change the

1571   crystal structure and the propensity for polymorphism. A recent survey of 232 simply substituted

1572   chalcones ((2*E*)-1,3-diphenylprop-2-en-1-ones)[385] had 170 different crystal packings, with the

1573   largest isomorphous group containing 15 compounds. Conversely, the digital design process could

1574   be used to foresee whether a given API was likely to be problematic to develop, and whether a

1575   multi-component form, or medically similar API should be considered.

1576   8.3   Utilization in Process Design

1577   Aspects of the vision of the digital design of crystallization processes are currently in use for

1578   drug development. CSP is used to understand the potential form landscape and inform the

1579   experimental form selection process. Morphology predictions are used to understand the effect

1580   solvent selection can have on the resultant crystal morphology, once a crystal form has been

1581   selected. However, both applications are employed in conjunction with experimental data to

1582   validate whether the potential energy surface and method are realistic enough – neither are used in

1583   the absence of experimental data, when only a molecular structure is available.

1584   Both tools could be used today *ab initio* to better understand and predict the risks associated

1585   with the development of a new API due to the complexity of the predicted solid form landscape

1586   and the expected crystal morphology landscape of each possible polymorph. The current state of

1587   the art of CSP, while limited by the accuracy of the potential energy surface and modeling of

1588   temperature effects, does usually identify the observed and stable polymorphs amongst other

1589   energetically competitive forms, as illustrated for both succinic acid[82] and for olanzapine (Figure

1590   5 and refs [73] and [79]). For olanzapine, as Figure 5 reveals, several forms that have not been observed

1591   experimentally are calculated to be more stable at ambient than those observed experimentally.

1592   However, Figure 4 shows that either force field used to model the molecular motions gives a very

1593   different lattice energy from the most advanced electronic structure methods, where the four

1594   known forms are the most stable. The general observation that CSP usually generates more

1595   thermodynamically plausible polymorphs than are found, even in state-of-the-art industrial

1596   screening (e.g. galunisertib[386]) and after landscape reduction (see Section 3.1.2), exemplifies the

1597   need for better primary nucleation models. It is conceivable that the rates of nucleation for the

1598   experimentally observed forms are faster than for the apparently thermodynamically competitive

1599   predicted yet not observed forms in all practically possible conditions.

1600   Table 1 illustrates the competition of the kinetics of nucleation and growth versus

1601   thermodynamics in predicting the form most likely to be observed when the free energy differences

1602   between competing forms are "small". In those cases when the more stable form exhibits faster

1603   kinetics for both nucleation and growth, the thermodynamic form should be experimentally

1604   observed. Similarly, when both the nucleation and growth kinetics of the thermodynamically

1605   stable form are much slower than the unstable form, the unstable form is more likely to be

1606   observed. It is not obvious which form would be expected when either the growth rate or the

1607   nucleation rate of the thermodynamically stable form is slow relative to the unstable form. For

1608   digital design to be able to predict the most likely form, better models are needed for nucleation to

1609   support the design of experiments to alter the nucleation rates and find the first sample of the new

1610   form (Section 3.3) or conclude that the putative structure was so kinetically hindered that it could

1611   never be crystallized. There are cases where the most stable form has eventually been found despite

1612   being slower to nucleate and grow than the apparently stable metastable form.[387, 388] However,

1613   once the solid form has been selected on the basis of experimental solid form screening with the

1614 aid of CSP, then the design of the crystallization process can proceed by ignoring the primary

1615 nucleation rate through seeding.

1616 Table 1: Experimentally expected form when the nucleation or growth rates of the stable form is

1617 fast/slow relative to the unstable forms.

Stable Form Nucleation Rate	Stable Form Growth Rate	Expected Form Observed
Fast	Slow	Unknown
Slow	Slow	Unstable Form
Fast	Fast	Stable Form
Slow	Fast	Unknown

1618

1619 More reasonable expectations of state-of-the-art molecular simulation at the present time or in

1620 the near future are to be able to:

1621   1. Approximate the stability order of a set of polymorphic forms as a function of
1622   temperature.

1623   2. Assess the risk of hydrate formation under real-world temperature and relative humidity
1624   conditions.

1625   3. Suggest the potential for further energetically competitive polymorphs and provide their
1626   structures, which may suggest an experimental route to their discovery.

1627   4. Predict the relative solubility of a given solute in different solvents.

1628   5. Predict the change in solubility with respect to temperature of a given solute in a given
1629   solvent (i.e., the slope of the van't Hoff plot (Figure 9)).

1630   6. Predict the morphology and the growth rates of the major facets for a set of polymorphic
1631   forms in a range of solvents.

1632   These are all important pieces of information needed during the development of a robust

1633   crystallization process. The methods we have described in this paper, based on the molecular

1634   mechanisms, using advanced MD sampling, etc are not yet commercially available.

1635   The inability to predict the most likely crystal form poses a conundrum for utilization of the

1636   digital design workflow as the predicted thermodynamically stable form may not be the form most

1637   likely to be observed and may not be the right basis to then use for *ab initio* conceptual process

1638   design. This suggests an alternative workflow for how the current state of the art in digital design

1639   can be used for conceptual crystallization process design – to carry forward all putative

1640   polymorphs within a reasonable free energy and hence solubility range (allowing for the errors in

1641   the simulation) and assess the crystal morphology landscape associated with each CSP-generated

1642   form. This workflow allows for a prediction of all crystal shapes possible and allows for resource

1643   and risk estimations to be made based on the complexity associated with the development of the

1644   crystallization process. As an example, if only equant crystal morphologies are predicted for all

1645   plausible polymorphs, then it can be anticipated that a smaller effort would be required to develop

1646   a crystallization process to provide a freely flowing powder. Conversely, if only needlelike

1647   crystals are predicted across all putative crystal forms, then it can be expected that more resources

1648   will be required to develop the crystalline API process to improve upon the crystal morphology

1649   through application of particle engineering techniques such as thermocycling.[389] For batch

1650   crystallization processes, it is possible to consider conceptual process design provided that

1651   solubility can be predicted. As this paper shows, *in silico* models provide a forward look at the

1652   effort needed to develop a new product, as the model predictions can provide an estimate of the

1653   complexity (undesired morphology, poor aqueous solubility, multiple polymorphs of similar free

1654   energies, poor organic solubility in acceptable process solvents) the solid form of the API might

1655   exhibit, and thus the impact (time and resource needs) on developing a solid oral dosage form.

1656   While it is not currently possible to fully realize the *ab initio* vision of digital design, it is possible

1657   to benefit from the workflow provided that small amounts of experimental data are available. For

1658   example, solubility in the full range of solvents can be broadly predicted from a small set of

1659   solubility measurements, and the absolute growth rate can be estimated from knowledge of a single

1660   absolute growth rate for only one face. Both examples represent an enormous reduction in effort

1661   compared with the experimental determination of the absolute growth rates of every face on the

1662   crystal surface (something which is rarely done, even in academic labs) or measurement of

1663   solubility in thousands of solvents or solvent mixtures.

1664   9   Conclusions

1665   An industrial perspective on the engineering of pharmaceutical materials in 2007[20] saw

1666   computational prediction as a future step in the route from active molecule to finished product. We

1667   have demonstrated through examples the use of simulation as part of the support to experimental

1668   techniques foreseen in a 2015 review of the future of pharmaceutical manufacturing sciences,[21]

1669   and illustrated much progress in some of the opportunities identified by the crystallization working

1670   group of the Enabling Technologies Consortium.[22]

1671   Computational methods are available to design a crystallization process from the chemical

1672   diagram (i.e. *a priori* without any experimental data) for all the main steps, apart from nucleation

1673   rates which are more important for polymorph discovery. However, which type and level of theory

1674   is accurate enough and practically affordable is very dependent on the specific molecule, its size,

1675   conformational flexibility, functional groups and particularly the nature of the competing

1676   polymorphs and hydrates that can affect the manufacture and stability of the product. Future efforts

1677   for utilizing CSP to aid polymorph discovery and minimizing the risks of the late emergence of

1678   more stable forms should focus on primary nucleation rates. A key requirement to the

1679   implementation of the digital design workflow is *a priori* solubility predictions but the exponential

1680   dependence of solubility on temperature and the free energy of solution (which can be greatly

1681   affected by anharmonic motions within the specific crystal) make this a particularly challenging

1682   property to compute. Morphologies can be computed, although this is more straightforward for

1683   centrosymmetric molecules or growth units, but absolute growth rates involve similar challenges

1684   to absolute solubilities.

1685   This paper shows that physics-based simulation methods are being actively developed for

1686   pharmaceuticals and are qualitatively and semi-quantitatively realistic enough to be useful. The

1687   ideal of using the accurate potential energy surfaces for CSP and all properties and a simulation

1688   method that is realistic for the dynamics, nucleation, and growth of pharmaceutical crystals, will

1689   eventually produce more reliable quantitative multi-scale modeling in the design of crystallization

1690   processes. In the meantime, such methods can considerably aid the design process when used

1691   either in conjunction with experimental data or for giving relative properties or key parameters in

1692   approximate models. The time is near when consideration of solid state properties for new

1693   chemical medicinal targets will be integrated into computational chemistry approaches for the

1694   identification of new drugs.

1695   10   Acknowledgements

11 References

(1) Deloitte. *Seize the digital momentum: Measuring the return from pharmaceutical innovation 2022*; Deloitte LLP, 2023. https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/life-sciences-health-care/deloitte-uk-seize-digital-momentum-rd-roi-2022.pdf (accessed 08 June 2023).
(2) *The Unbearable Cost of Drug Development: Deloitte Report Shows 15% Jump in R&D to $2.3 Billion*. 2023. https://www.genengnews.com/gen-edge/the-unbearable-cost-of-drug-development-deloitte-report-shows-15-jump-in-rd-to-2-3-billion/ (accessed 2023 08 June 2023).
(3) DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* **2016**, *47*, 20-33.
(4) Ooms, F. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. *Curr. Med. Chem.* **2000**, *7* (2), 141-158.
(5) Andricopulo, A. D.; Salum, L. B.; Abraham, D. J. Structure-based drug design strategies in medicinal chemistry. *Curr. Top. Med. Chem. (Sharjah, United Arab Emirates)* **2009**, *9* (9), 771-790.
(6) Mangiatordi, G. F.; Carotti, A.; Novellino, E.; Nicolotti, O. A round trip from medicinal chemistry to predictive toxicology. *Methods Mol. Biol. (N. Y., NY, U. S.)* **2016**, *1425*, 461-473.
(7) Abel, R.; Wang, L.; Harder, E. D.; Berne, B. J.; Friesner, R. A. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Accounts of Chemical Research* **2017**, *50* (7), 1625-1632.
(8) Kuhn, B.; Tichý, M.; Wang, L.; Robinson, S.; Martin, R. E.; Kuglstatter, A.; Benz, J.; Giroud, M.; Schirmeister, T.; Abel, R.; et al. Prospective Evaluation of Free Energy Calculations for the Prioritization of Cathepsin L Inhibitors. *Journal of Medicinal Chemistry* **2017**, *60* (6), 2485-2497.
(9) Soltani, S.; Hallaj-Nezhadi, S.; Rashidi, M. R. A comprehensive review of in silico approaches for the prediction and modulation of aldehyde oxidase-mediated drug metabolism: The current features, challenges and future perspectives. *European Journal of Medicinal Chemistry* **2021**, *222*, 113559.
(10) Richard, A. M. Future of Toxicology-Predictive Toxicology: An Expanded View of "Chemical Toxicity". *Chem. Res. Toxicol.* **2006**, *19* (10), 1257-1262.
(11) Valerio, L. G.; Cross, K. P. Characterization and validation of an in silico toxicology model to predict the mutagenic potential of drug impurities*. *Toxicol. Appl. Pharmacol.* **2012**, *260* (3), 209-221.
(12) Silverman, R. B. Chapter 2 - Drug Discovery, Design, and Development. In *The Organic Chemistry of Drug Design and Drug Action (Second Edition)*, Silverman, R. B. Ed.; Academic Press, 2004; pp 7-120.
(13) Amidon, G. L.; Lennernäs, H.; Shah, V. P.; Crison, J. R. A Theoretical Basis for a Biopharmaceutic Drug Classification: The Correlation of in Vitro Drug Product Dissolution and in Vivo Bioavailability. *Pharmaceutical Research* **1995**, *12* (3), 413-420.

1741 (14) Oh, D.-M.; Curl, R. L.; Amidon, G. L. Estimating the Fraction Dose Absorbed from
1742 Suspensions of Poorly Soluble Compounds in Humans: A Mathematical Model. *Pharmaceutical*
1743 *Research* **1993**, *10* (2), 264-270.
1744 (15) Yu, L. X.; Lipka, E.; Crison, J. R.; Amidon, G. L. Transport approaches to the
1745 biopharmaceutical design of oral drug delivery systems: prediction of intestinal absorption.
1746 *Advanced Drug Delivery Reviews* **1996**, *19* (3), 359-376.
1747 (16) Kayala, M. A.; Baldi, P. ReactionPredictor: Prediction of Complex Chemical Reactions at
1748 the Mechanistic Level Using Machine Learning. *Journal of Chemical Information and Modeling*
1749 **2012**, *52* (10), 2526-2540.
1750 (17) Liu, B.; Ramsundar, B.; Kawthekar, P.; Shi, J.; Gomes, J.; Luu Nguyen, Q.; Ho, S.; Sloane,
1751 J.; Wender, P.; Pande, V. Retrosynthetic Reaction Prediction Using Neural Sequence-to-
1752 Sequence Models. *ACS Central Science* **2017**, *3* (10), 1103-1113.
1753 (18) Schwaller, P.; Petraglia, R.; Zullo, V.; Nair, V. H.; Haeuselmann, R. A.; Pisoni, R.; Bekas,
1754 C.; Iuliano, A.; Laino, T. Predicting retrosynthetic pathways using transformer-based models and
1755 a hyper-graph exploration strategy. *Chemical Science* **2020**, *11* (12), 3316-3325.
1756 (19) Segler, M. H. S.; Preuss, M.; Waller, M. P. Planning chemical syntheses with deep neural
1757 networks and symbolic AI. *Nature* **2018**, *555* (7698), 604-610.
1758 (20) Chow, K.; Tong, H. H. Y.; Lum, S.; Chow, A. H. L. Engineering of pharmaceutical
1759 materials: An industrial perspective. *Journal of Pharmaceutical Sciences* **2008**, *97* (8), 2855-
1760 2877.
1761 (21) Rantanen, J.; Khinast, J. The Future of Pharmaceutical Manufacturing Sciences. *Journal of*
1762 *Pharmaceutical Sciences* **2015**, *104* (11), 3612-3638.
1763 (22) Cote, A.; Erdemir, D.; Girard, K. P.; Green, D. A.; Lovette, M. A.; Sirota, E.; Nere, N. K.
1764 Perspectives on the Current State, Challenges, and Opportunities in Pharmaceutical
1765 Crystallization Process Development. *Crystal Growth & Design* **2020**, *20* (12), 7568-7581.
1766 (23) Abramov, Y. A.; Sun, G. X.; Zeng, Q. Emerging Landscape of Computational Modeling in
1767 Pharmaceutical Development. *Journal of Chemical Information and Modeling* **2022**, *62* (5),
1768 1160-1171.
1769 (24) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg,
1770 J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; et al. Report on the sixth blind test of
1771 organic crystal structure prediction methods. *Acta Crystallographica Section B* **2016**, *72* (4),
1772 439-459.
1773 (25) Llinas, A.; Oprisiu, I.; Avdeef, A. Findings of the Second Challenge to Predict Aqueous
1774 Solubility. *Journal of Chemical Information and Modeling* **2020**, *60* (10), 4791-4803.
1775 (26) Price, S. L. Progress in understanding crystallisation: a personal perspective. *Faraday*
1776 *Discussions* **2022**, *235*, 569-581.
1777 (27) Lovette, M. A.; Albrecht, J.; Ananthula, R. S.; Ricci, F.; Sangodkar, R.; Shah, M. S.;
1778 Tomasi, S. Evaluation of Predictive Solubility Models in Pharmaceutical Process
1779 Development─an Enabling Technologies Consortium Collaboration. *Crystal Growth & Design*
1780 **2022**, *22* (9), 5239-5263.
1781 (28) Feeder, N.; Pidcock, E.; Reilly, A. M.; Sadiq, G.; Doherty, C. L.; Back, K. R.; Meenan, P.;
1782 Docherty, R. The integration of solid-form informatics into solid-form selection. *Journal of*
1783 *Pharmacy and Pharmacology* **2015**, *67* (6), 857-868.
1784 (29) Reutzel-Edens, S. M.; Bhardwaj, R. M. Crystal forms in pharmaceutical applications:
1785 olanzapine, a gift to crystal chemistry that keeps on giving. *Iucrj* **2020**, *7*, 955-964.

1786 (30) Boye, K.; Ross, M.; Mody, R.; Konig, M.; Gelhorn, H. Patients' preferences for once-daily
1787 oral versus once-weekly injectable diabetes medications: The REVISE study. *Diabetes, Obesity*
1788 *and Metabolism* **2021**, *23* (2), 508-519.
1789 (31) Iacocca, R. G.; Burcham, C. L.; Hilden, L. R. Particle engineering: A strategy for
1790 establishing drug substance physical property specifications during small molecule development.
1791 *Journal of Pharmaceutical Sciences* **2010**, *99* (1), 51-75.
1792 (32) Abramov, Y. A. Rational Solvent Selection for Pharmaceutical Impurity Purge. *Crystal*
1793 *Growth & Design* **2018**, *18* (2), 1208-1214.
1794 (33) Abramov, Y. A.; Zelellow, A.; Chen, C.-y.; Wang, J.; Sekharan, S. Novel Computational
1795 Approach to Guide Impurities Rejection by Crystallization: A Case Study of MRTX849
1796 Impurities. *Crystal Growth & Design* **2022**, *22* (12), 6844-6848.
1797 (34) Burcham, C. L.; Florence, A. J.; Johnson, M. D. Continuous Manufacturing in
1798 Pharmaceutical Process Development and Manufacturing. *Annual Review of Chemical and*
1799 *Biomolecular Engineering* **2018**, *9* (1), 253-281.
1800 (35) Johnson, M. C. R. Particle size distribution of the active ingredient for solid dosage forms of
1801 low dosage. *Pharmaceutica Acta Helvetiae* **1972**, *47*, 546-559.
1802 (36) Yalkowsky, S. H.; Bolton, S. Particle Size and Content Uniformity. *Pharmaceutical*
1803 *Research* **1990**, *7* (9), 962-966.
1804 (37) Rohrs, B. R.; Amidon, G. E.; Meury, R. H.; Secreast, P. J.; King, H. M.; Skoug, C. J.
1805 Particle Size Limits to Meet USP Content Uniformity Criteria for Tablets and Capsules. *Journal*
1806 *of Pharmaceutical Sciences* **2006**, *95* (5), 1049-1059.
1807 (38) Bauer, J.; Spanton, S.; Henry, R.; Quick, J.; Dziki, W.; Porter, W.; Morris, J. Ritonavir: An
1808 extraordinary example of conformational polymorphism. *Pharmaceutical Research* **2001**, *18* (6),
1809 859-866.
1810 (39) Bucar, D. K.; Lancaster, R. W.; Bernstein, J. Disappearing Polymorphs Revisited.
1811 *Angewandte Chemie-International Edition* **2015**, *54* (24), 6972-6993.
1812 (40) Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.;
1813 Spanton, S.; Dziki, W.; Porter, W.; et al. Dealing with the impact of ritonavir polymorphs on the
1814 late stages of bulk drug process development. *Organic Process Research & Development* **2000**,
1815 *4* (5), 413-417.
1816 (41) Morissette, S. L.; Soukasene, S.; Levinson, D.; Cima, M. J.; Almarsson, O. Eludication of
1817 crystal form diversity of the HIV protease inhibitor ritonavir by high-throughput crystallisation.
1818 *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100* (5),
1819 2180-2184.
1820 (42) Yamaguchi, S.; Kaneko, M.; Narukawa, M. Approval success rates of drug candidates based
1821 on target, action, modality, application, and their combinations. *Clinical and Translational*
1822 *Science* **2021**, *14* (3), 1113-1122.
1823 (43) Bernstein, J. *Polymorphism in Molecular Crystals*; Clarendon Press, 2020.
1824 (44) Reutzel-Edens, S. M.; Braun, D. E.; Newman, A. W. Hygroscopicity and Hydrates in
1825 Pharmaceutical Solids. In *Polymorphism in the Pharmaceutical Industry*, 2018; pp 159-188.
1826 (45) Price, S. L.; Reutzel-Edens, S. M. The potential of computed crystal energy landscapes to
1827 aid solid-form development. *Drug Discovery Today* **2016**, *21* (6), 912-923.
1828 (46) Price, S. L. Is zeroth order crystal structure prediction (CSP_0) coming to maturity? What
1829 should we aim for in an ideal crystal structure prediction code? *Faraday Discussions* **2018**, *211*,
1830 9-30.

1831 (47) Nyman, J.; Reutzel-Edens, S. M. Crystal structure prediction is changing from basic science
1832 to applied technology. *Faraday Discussions* **2018**, *211*, 459-476.
1833 (48) Belenguer, A. M.; Lampronti, G. I.; Cruz-Cabeza, A. J.; Hunter, C. A.; Sanders, J. K. M.
1834 Solvation and surface effects on polymorph stabilities at the nanoscale. *Chemical Science* **2016**,
1835 *7* (11), 6617-6627.
1836 (49) Ward, M. D. Perils of Polymorphism: Size Matters. *Israel Journal of Chemistry* **2017**, *57*
1837 (1-2), 82-91.
1838 (50) Firaha, D.; Liu, Y. M.; van de Streek, J.; Sasikumar, K.; Dietrich, H.; Helfferich, J.; Aerts,
1839 L.; Braun, D. E.; Broo, A.; DiPasquale, A. G.; et al. Predicting crystal form stability under real-
1840 world conditions. *Nature* **2023**, *623* (7986), 324-328.
1841 (51) Braun, D. E.; McMahon, J. A.; Bhardwaj, R. M.; Nyman, J.; Neumann, M. A.; van de
1842 Streek, J.; Reutzel-Edens, S. M. Inconvenient Truths about Solid Form Landscapes Revealed in
1843 the Polymorphs and Hydrates of Gandotinib. *Crystal Growth & Design* **2019**, *19* (5), 2947-2962.
1844 (52) Price, S. L. Why don't we find more polymorphs? *Acta Crystallographica Section B:*
1845 *Structural Science, Crystal Engineering and Materials* **2013**, *69*, 313-328.
1846 (53) Sosso, G. C.; Chen, J.; Cox, S. J.; Fitzner, M.; Pedevilla, P.; Zen, A.; Michaelides, A.
1847 Crystal Nucleation in Liquids: Open Questions and Future Challenges in Molecular Dynamics
1848 Simulations. *Chemical Reviews* **2016**, *116* (12), 7078-7116.
1849 (54) Woollam, G. R.; Neumann, M. A.; Wagner, T.; Davey, R. J. The importance of
1850 configurational disorder in crystal structure prediction: the case of loratadine. *Faraday*
1851 *Discussions* **2018**, *211*, 209-234.
1852 (55) Neumann, M. A. Tailor-made force fields for crystal-structure prediction. *Journal of*
1853 *Physical Chemistry B* **2008**, *112* (32), 9810-9829.
1854 (56) Price, S. L. Control and prediction of the organic solid state: a challenge to theory and
1855 experiment. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering*
1856 *Science* **2018**, *474* (2217), 20180351.
1857 (57) Day, G. M. Current approaches to predicting molecular organic crystal structures.
1858 *Crystallography Reviews* **2011**, *17* (1), 3-52.
1859 (58) Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.;
1860 Day, G. M. Modelling Organic Crystal Structures using Distributed Multipole and Polarizability-
1861 Based Model Intermolecular Potentials. *Physical Chemistry Chemical Physics* **2010**, *12* (30),
1862 8478-8490.
1863 (59) Day, G. M.; Motherwell, W. D. S.; Jones, W. Beyond the isotropic atom model in crystal
1864 structure prediction of rigid molecules: Atomic multipoles versus point charges. *Crystal Growth*
1865 *& Design* **2005**, *5* (3), 1023-1033.
1866 (60) Price, S. L. Computed crystal energy landscapes for understanding and predicting organic
1867 crystal structures and polymorphism. *Accounts of Chemical Research* **2009**, *42* (1), 117-126.
1868 (61) Aina, A. A.; Misquitta, A. J.; Price, S. L. A non-empirical intermolecular force-field for
1869 trinitrobenzene and its application in crystal structure prediction. *The Journal of Chemical*
1870 *Physics* **2021**, *154* (9), 094123.
1871 (62) Greenwell, C.; McKinley, J. L.; Zhang, P. Y.; Zeng, Q.; Sun, G. X.; Li, B. C.; Wen, S. H.;
1872 Beran, G. J. O. Overcoming the difficulties of predicting conformational polymorph energetics in
1873 molecular crystals via correlated wavefunction methods. *Chemical Science* **2020**, *11* (8), 2200-
1874 2214.
1875 (63) Bowskill, D. H.; Sugden, I. J.; Konstantinopoulos, S.; Adjiman, C. S.; Pantelides, C. C.;
1876 Doherty, M. F.; Segalman, R. A. Crystal Structure Prediction Methods for Organic Molecules:

State of the Art. *Annual Review of Chemical and Biomolecular Engineering, Vol 12, 2021* **2021**, *12*, 593-623.

(64) Cooper, T. G.; Hejczyk, K. E.; Jones, W.; Day, G. M. Molecular Polarization Effects on the Relative Energies of the Real and Putative Crystal Structures of Valine. *Journal of Chemical Theory and Computation* **2008**, *4* (10), 1795-1805.

(65) Hoja, J.; Ko, H.-Y.; Neumann, M. A.; Car, R.; DiStasio, R. A.; Tkatchenko, A. Reliable and practical computational description of molecular crystal polymorphs. *Science Advances* **2019**, *5* (1), eaau-3338.

(66) LeBlanc, L. M.; Otero-de-la-Roza, A.; Johnson, E. R. Composite and Low-Cost Approaches for Molecular Crystal Structure Prediction. *Journal of Chemical Theory and Computation* **2018**, *14* (4), 2265-2276.

(67) Beran, G. J. O. Modeling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chemical Reviews* **2016**, *116* (9), 5567-5613.

(68) Braun, D. E.; Ardid-Candel, M.; D'Oria, E.; Karamertzanis, P. G.; Arlin, J. B.; Florence, A. J.; Jones, A. G.; Price, S. L. Racemic Naproxen: A Multidisciplinary Structural and Thermodynamic Comparison with the Enantiopure Form. *Crystal Growth & Design* **2011**, *11* (12), 5659-5669.

(69) Raiteri, P.; Martonak, R.; Parrinello, M. Exploring Polymorphism: The Case of Benzene. *Angewandte Chemie-International Edition* **2005**, *44*, 3769-3773.

(70) Francia, N. F.; Price, L. S.; Salvalaglio, M. Reducing crystal structure overprediction of ibuprofen with large scale molecular dynamics simulations. *CrystEngComm* **2021**, *23* (33), 5575-5584.

(71) Francia, N. F.; Price, L. S.; Nyman, J.; Price, S. L.; Salvalaglio, M. Systematic Finite-Temperature Reduction of Crystal Energy Landscapes. *Crystal Growth & Design* **2020**, *20*, 6847-6862.

(72) Sugden, I. J.; Francia, N. F.; Jensen, T.; Adjiman, C. S.; Salvalaglio, M. Rationalising the difference in crystallisability of two sulflowers using efficient in silico methods. *Crystengcomm* **2022**, *24* (39), 6830-6838.

(73) Bhardwaj, R. M.; Price, L. S.; Price, S. L.; Reutzel-Edens, S. M.; Miller, G. J.; Oswald, I. D. H.; Johnston, B.; Florence, A. J. Exploring the Experimental and Computed Crystal Energy Landscape of Olanzapine. *Crystal Growth & Design* **2013**, *13* (4), 1602-1617.

(74) Price, L. S.; McMahon, J. A.; Lingireddy, S. R.; Lau, S. F.; Diseroad, B. A.; Price, S. L.; Reutzel-Edens, S. M. A molecular picture of the problems in ensuring structural purity of tazofelone. *Journal of Molecular Structure* **2014**, *1078*, 26-42.

(75) Bond, A. D.; Boese, R.; Desiraju, G. R. On the polymorphism of aspirin: Crystalline aspirin as intergrowths of two "polymorphic" domains. *Angewandte Chemie-International Edition* **2007**, *46* (4), 618-622.

(76) Warzecha, M.; Guo, R.; Bhardwaj, R. M.; Reutzel-Edens, S. M.; Price, S. L.; Lamprou, D. A.; Florence, A. J. Direct Observation of Templated Two-Step Nucleation Mechanism during Olanzapine Hydrate Formation. *Crystal Growth & Design* **2017**, *17* (12), 6382–6393.

(77) Warzecha, M.; Safari, M. S.; Florence, A. J.; Vekilov, P. G. Mesoscopic Solute-Rich Clusters in Olanzapine Solutions. *Crystal Growth & Design* **2017**, *17* (12), 6668-6676.

(78) Warzecha, M.; Verma, L.; Johnston, B. F.; Palmer, J. C.; Florence, A. J.; Vekilov, P. G. Olanzapine crystal symmetry originates in preformed centrosymmetric solute dimers. *Nature Chemistry* **2020**, *12* (10), 914-+.

1922 (79) Askin, S.; Cockcroft, J. K.; Price, L. S.; Goncalves, A. D.; Zhao, M.; Tocher, D. A.;
1923 Williams, G. R.; Gaisford, S.; Craig, D. Q. M. Olanzapine Form IV: Discovery of a New
1924 Polymorphic Form Enabled by Computed Crystal Energy Landscapes. *Crystal Growth & Design*
1925 **2019**, *19* (5), 2751-2757.
1926 (80) LeBlanc, L. M.; Johnson, E. R. Crystal-energy landscapes of active pharmaceutical
1927 ingredients using composite approaches. *CrystEngComm* **2019**, *21* (40), 5995-6009.
1928 (81) Askin, S.; Goncalves, A. D.; Zhao, M.; Williams, G. R.; Gaisford, S.; Craig, D. Q. M. A
1929 Simultaneous Differential Scanning Calorimetry-X-ray Diffraction Study of Olanzapine
1930 Crystallization from Amorphous Solid Dispersions. *Molecular Pharmaceutics* **2020**, *17* (11),
1931 4364-4374.
1932 (82) Lucaioli, P.; Nauha, E.; Gimondi, I.; Price, L. S.; Guo, R.; Iuzzolino, L.; Singh, I.;
1933 Salvalaglio, M.; Price, S. L.; Blagden, N. Serendipitous isolation of a disappearing
1934 conformational polymorph of succinic acid challenges computational polymorph prediction.
1935 *CrystEngComm* **2018**, *20* (28), 3971-3977.
1936 (83) van Eijck, B. P.; Kroon-Batenburg, L. M. J.; Kroon, J. Energy minimisation and Molecular
1937 Dynamics calculations for molecular crystals. In *Theoretical Aspects and Computer Modeling of*
1938 *the Molecular Solid State*, Gavezzotti, A. Ed.; John Wiley & Sons, 1997; pp 99-146.
1939 (84) Gavezzotti, A. A molecular dynamics test of the different stability of crystal polymorphs
1940 under thermal strain. *Journal of the American Chemical Society* **2000**, *122* (43), 10724-10725.
1941 (85) Yang, S.; Day, G. M. Global analysis of the energy landscapes of molecular crystal
1942 structures by applying the threshold algorithm. *Communications Chemistry* **2022**, *5* (1), 86.
1943 (86) Nyman, J.; Day, G. M. Modelling temperature-dependent properties of polymorphic organic
1944 molecular crystals. *Physical Chemistry Chemical Physics* **2016**, *18* (45), 31132-31143.
1945 (87) Buchholz, H. K.; Hylton, R. K.; Brandenburg, J. G.; Seidel-Morgenstern, A.; Lorenz, H.;
1946 Stein, M.; Price, S. L. Thermochemistry of Racemic and Enantiopure Organic Crystals for
1947 Predicting Enantiomer Separation. *Crystal Growth & Design* **2017**, *17* (9), 4676-4686.
1948 (88) Reilly, A. M.; Tkatchenko, A. Role of Dispersion Interactions in the Polymorphism and
1949 Entropic Stabilization of the Aspirin Crystal. *Physical Review Letters* **2014**, *113* (5), 055701.
1950 (89) Srirambhatla, V. K.; Guo, R.; Dawson, D. M.; Price, S. L.; Florence, A. J. Reversible, Two-
1951 Step Single-Crystal to Single-Crystal Phase Transitions between Desloratadine Forms I, II, and
1952 III. *Crystal Growth & Design* **2020**, *20* (3), 1800-1810.
1953 (90) Potticary, J.; Hall, C. L.; Guo, R.; Price, S. L.; Hall, S. R. On the Application of Strong
1954 Magnetic Fields during Organic Crystal Growth. *Crystal Growth & Design* **2021**, *21* (11), 6254-
1955 6265.
1956 (91) Luo, H. Y.; Hao, X.; Gong, Y. Q.; Zhou, J. H.; He, X.; Li, J. J. Rational Crystal Polymorph
1957 Design of Olanzapine. *Crystal Growth & Design* **2019**, *19* (4), 2388-2395.
1958 (92) Tang, J. Q.; Han, Y. Q.; Ali, I.; Luo, H. Y.; Nowak, A.; Li, J. J. Stability and phase
1959 transition investigation of olanzapine polymorphs. *Chemical Physics Letters* **2021**, *767*, 138384.
1960 (93) Brandenburg, J. G.; Potticary, J.; Sparkes, H. A.; Price, S. L.; Hall, S. R. Thermal Expansion
1961 of Carbamazepine: Systematic Crystallographic Measurements Challenge Quantum Chemical
1962 Calculations. *Journal of Physical Chemistry Letters* **2017**, *8* (17), 4319-4324.
1963 (94) McKinley, J. L.; Beran, G. J. O. Identifying pragmatic quasi-harmonic electronic structure
1964 approaches for modeling molecular crystal thermal expansion. *Faraday Discussions* **2018**, *211*,
1965 181-207.

1966 (95) Souvatzis, P.; Eriksson, O.; Katsnelson, M. I.; Rudin, S. P. Entropy Driven Stabilization of
1967 Energetically Unstable Crystal Structures Explained from First Principles Theory. *Physical*
1968 *Review Letters* **2008**, *100* (9), 095901.
1969 (96) Hellman, O.; Abrikosov, I. A.; Simak, S. I. Lattice dynamics of anharmonic solids from first
1970 principles. *Physical Review B* **2011**, *84* (18), 180301.
1971 (97) Kapil, V.; Engel, E. A. A complete description of thermodynamic stabilities of molecular
1972 crystals. *Proceedings of the National Academy of Sciences* **2022**, *119* (6), e2111769119.
1973 (98) Hill, T. L. *An Introduction to Statistical Tharmodynamics*; Addison-Wesley Publishing
1974 Company, Inc.; Courier Corporation, 1960, 1986.
1975 (99) Tuckerman, M. E. *Statistical Mechanics: Theory and Molecular Simulation*; Oxford
1976 University Press, 2010.
1977 (100) Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to*
1978 *applications*; Elsevier, 2001.
1979 (101) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-
1980 energy estimation: Umbrella sampling. *Journal of Computational Physics* **1977**, *23* (2), 187-199.
1981 (102) Roux, B. The calculation of the potential of mean force using computer simulations.
1982 *Computer Physics Communications* **1995**, *91* (1), 275-282.
1983 (103) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the National*
1984 *Academy of Sciences of the United States of America* **2002**, *99* (20), 12562-12566.
1985 (104) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly
1986 converging and tunable free-energy method. *Physical Review Letters* **2008**, *100* (2), 020603.
1987 (105) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *WIREs Computational*
1988 *Molecular Science* **2011**, *1* (5), 826-843.
1989 (106) Kästner, J. Umbrella sampling. *WIREs Computational Molecular Science* **2011**, *1* (6), 932-
1990 942.
1991 (107) Kästner, J.; Thiel, W. Bridging the gap between thermodynamic integration and umbrella
1992 sampling provides a novel analysis method: "Umbrella integration". *The Journal of Chemical*
1993 *Physics* **2005**, *123* (14), 144104.
1994 (108) Marinova, V.; Salvalaglio, M. Time-independent free energies from metadynamics via
1995 mean force integration. *The Journal of Chemical Physics* **2019**, *151* (16), 164115.
1996 (109) Giberti, F.; Cheng, B.; Tribello, G. A.; Ceriotti, M. Iterative Unbiasing of Quasi-
1997 Equilibrium Sampling. *Journal of Chemical Theory and Computation* **2020**, *16* (1), 100-107.
1998 (110) Tiwary, P.; Parrinello, M. A Time-Independent Free Energy Estimator for Metadynamics.
1999 *The Journal of Physical Chemistry B* **2015**, *119* (3), 736-742.
2000 (111) Bonomi, M.; Barducci, A.; Parrinello, M. Reconstructing the equilibrium Boltzmann
2001 distribution from well-tempered metadynamics. *Journal of Computational Chemistry* **2009**, *30*
2002 (11), 1615-1621.
2003 (112) Cuendet, M. A.; Tuckerman, M. E. Free Energy Reconstruction from Metadynamics or
2004 Adiabatic Free Energy Dynamics Simulations. *Journal of Chemical Theory and Computation*
2005 **2014**, *10* (8), 2975-2986.
2006 (113) Rosso, L.; Tuckerman, M. E. An Adiabatic Molecular Dynamics Method for the
2007 Calculation of Free Energy Profiles. *Molecular Simulation* **2002**, *28* (1-2), 91-112.
2008 (114) Schneider, E.; Vogt, L.; Tuckerman, M. E. Exploring polymorphism of benzene and
2009 naphthalene with free energy based enhanced molecular dynamics. *Acta Crystallographica*
2010 *Section B-Structural Science Crystal Engineering and Materials* **2016**, *72*, 542-550.

2011 (115) Abrams, J. B.; Tuckerman, M. E. Efficient and Direct Generation of Multidimensional
2012 Free Energy Surfaces via Adiabatic Dynamics without Coordinate Transformations. *The Journal*
2013 *of Physical Chemistry B* **2008**, *112* (49), 15742-15757.
2014 (116) Yu, T.-Q.; Tuckerman, M. E. Temperature-Accelerated Method for Exploring
2015 Polymorphism in Molecular Crystals Based on Free Energy. *Physical Review Letters* **2011**, *107*
2016 (1), 015701.
2017 (117) Yu, T.-Q.; Chen, P.-Y.; Chen, M.; Samanta, A.; Vanden-Eijnden, E.; Tuckerman, M.
2018 Order-parameter-aided temperature-accelerated sampling for the exploration of crystal
2019 polymorphism and solid-liquid phase transitions. *The Journal of Chemical Physics* **2014**, *140*
2020 (21), 214109.
2021 (118) Giberti, F.; Salvalaglio, M.; Mazzotti, M.; Parrinello, M. Insight into the nucleation of urea
2022 crystals from the melt. *Chemical Engineering Science* **2015**, *121*, 51-59.
2023 (119) Giberti, F.; Salvalaglio, M.; Parrinello, M. Metadynamics studies of crystal nucleation.
2024 *IUCrJ* **2015**, *2* (2), 256-266.
2025 (120) Gimondi, I.; Salvalaglio, M. CO2 packing polymorphism under pressure: Mechanism and
2026 thermodynamics of the I-III polymorphic transition. *The Journal of Chemical Physics* **2017**, *147*
2027 (11), 114502.
2028 (121) Piaggi, P. M.; Parrinello, M. Predicting polymorphism in molecular crystals using
2029 orientational entropy. *Proceedings of the National Academy of Sciences* **2018**, *115* (41), 10251-
2030 10256.
2031 (122) Piaggi, P. M.; Valsson, O.; Parrinello, M. Enhancing Entropy and Enthalpy Fluctuations to
2032 Drive Crystallization in Atomistic Simulations. *Physical Review Letters* **2017**, *119* (1), 015701.
2033 (123) Song, H. X.; Vogt-Maranto, L.; Wiscons, R.; Matzger, A. J.; Tuckerman, M. E. Generating
2034 Cocrystal Polymorphs with Information Entropy Driven by Molecular Dynamics-Based
2035 Enhanced Sampling. *Journal of Physical Chemistry Letters* **2020**, *11* (22), 9751-9758.
2036 (124) Metz, M. P.; Shahbaz, M.; Song, H. X.; Vogt-Maranto, L.; Tuckerman, M. E.; Szalewicz,
2037 K. Crystal Structure Predictions for 4-Amino-2,3,6-trinitrophenol Using a Tailor-Made First-
2038 Principles-Based Force Field. *Crystal Growth & Design* **2022**, *22* (2), 1182-1195.
2039 (125) Gobbo, G.; Bellucci, M. A.; Tribello, G. A.; Ciccotti, G.; Trout, B. L. Nucleation of
2040 Molecular Crystals Driven by Relative Information Entropy. *Journal of Chemical Theory and*
2041 *Computation* **2018**, *14* (2), 959-972.
2042 (126) Frenkel, D.; Ladd, A. J. C. New Monte Carlo method to compute the free energy of
2043 arbitrary solids. Application to the fcc and hcp phases of hard spheres. *The Journal of Chemical*
2044 *Physics* **1984**, *81* (7), 3188-3193.
2045 (127) Vega, C.; Noya, E. G. Revisiting the Frenkel-Ladd method to compute the free energy of
2046 solids: The Einstein molecule approach. *The Journal of Chemical Physics* **2007**, *127* (15),
2047 154113.
2048 (128) Sweatman, M. B.; Atamas, A. A.; Leyssale, J.-M. The self-referential method combined
2049 with thermodynamic integration. *The Journal of Chemical Physics* **2008**, *128* (6), 064102.
2050 (129) Frenkel, D.; Smit, B. *Understanding Molecular Simulations: from Algorithms to*
2051 *Applications*; Elsevier, 2002.
2052 (130) Monson, P. A.; Kofke, D. A. Solid-Fluid Equilibrium: Insights from Simple Molecular
2053 Models. *Advances in Chemical Physics* **2000**, 113-179.
2054 (131) Li, L.; Totton, T.; Frenkel, D. Computational methodology for solubility prediction:
2055 Application to the sparingly soluble solutes. *The Journal of Chemical Physics* **2017**, *146* (21),
2056 214110.

(132) Aragones, J. L.; Valeriani, C.; Vega, C. Note: Free energy calculations for atomic solids through the Einstein crystal/molecule methodology using GROMACS and LAMMPS. *The Journal of Chemical Physics* **2012**, *137* (14), 146101.

(133) Aragones, J. L.; Noya, E. G.; Valeriani, C.; Vega, C. Free energy calculations for molecular solids using GROMACS. *The Journal of Chemical Physics* **2013**, *139* (3), 034104.

(134) Addula, R. K. R.; Punnathanam, S. N. Calculation of excess free energy of molecular solids comprised of flexible molecules using Einstein molecule method. *Molecular Simulation* **2018**, *44* (10), 781-788.

(135) Bruce, A. D.; Wilding, N. B.; Ackland, G. J. Free Energy of Crystalline Solids: A Lattice-Switch Monte Carlo Method. *Physical Review Letters* **1997**, *79* (16), 3002-3005.

(136) Bruce, A. D.; Jackson, A. N.; Ackland, G. J.; Wilding, N. B. Lattice-switch Monte Carlo method. *Physical Review E* **2000**, *61* (1), 906-919.

(137) Wilms, D.; Wilding, N. B.; Binder, K. Transitions between imperfectly ordered crystalline structures: A phase switch Monte Carlo study. *Physical Review E* **2012**, *85* (5), 056703.

(138) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *The Journal of Chemical Physics* **1954**, *22* (8), 1420-1426.

(139) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *Journal of Computational Physics* **1976**, *22* (2), 245-268.

(140) Kamat, K.; Peters, B. Diabat Interpolation for Polymorph Free-Energy Differences. *Journal of Physical Chemistry Letters* **2017**, *8* (3), 655-660.

(141) Kamat, K.; Peters, B. Gibbs free-energy differences between polymorphs via a diabat approach. *The Journal of Chemical Physics* **2018**, *149* (21), 214106.

(142) Kamat, K.; Guo, R.; Reutzel-Edens, S. M.; Price, S. L.; Peters, B. Diabat method for polymorph free energies: Extension to molecular crystals. *The Journal of Chemical Physics* **2020**, *153* (24), 244105.

(143) Cruz-Cabeza, A. J.; Feeder, N.; Davey, R. J. Open questions in organic crystal polymorphism. *Communications Chemistry* **2020**, *3* (1).

(144) Brittain, H. G. *Polymorphism in Pharmaceutical Solids*; Informa Healthcare, 2016.

(145) Neumann, M. A.; van de Streek, J. How many ritonavir cases are there still out there? *Faraday Discussions* **2018**, *211*, 441-458.

(146) Neumann, M. A.; de Streek, J. V.; Fabbiani, F. P. A.; Hidber, P.; Grassmann, O. Combined crystal structure prediction and high-pressure crystallization in rational pharmaceutical polymorph screening. *Nature Communications* **2015**, *6*, 7793.

(147) Oswald, I. D. H.; Chataigner, I.; Elphick, S.; Fabbiani, F. P. A.; Lennie, A. R.; Maddaluno, J.; Marshall, W. G.; Prior, T. J.; Pulham, C. R.; Smith, R. I. Putting pressure on elusive polymorphs and solvates. *CrystEngComm* **2009**, *11* (2), 359-366.

(148) Day, G. M.; Cooper, A. I. Energy-Structure-Function Maps: Cartography for Materials Discovery. *Advanced Materials* **2018**, *30* (37), 1704944.

(149) Greenaway, R. L.; Jelfs, K. E. Integrating Computational and Experimental Workflows for Accelerated Organic Materials Discovery. *Advanced Materials* **2021**, *33* (11), 2004831.

(150) McArdle, P.; Erxleben, A. Sublimation - a green route to new solid-state forms. *CrystEngComm* **2021**, *23*, 5965-5975.

(151) Srirambhatla, V. K.; Guo, R.; Price, S. L.; Florence, A. J. Isomorphous template induced crystallisation: a robust method for the targeted crystallisation of computationally predicted metastable polymorphs. *Chemical Communications* **2016**, *52*, 7384-7386.

(152) Case, D. H.; Srirambhatla, V. K.; Guo, R.; Watson, R. E.; Price, L. S.; Polyzois, H.; Cockcroft, J. K.; Florence, A. J.; Tocher, D. A.; Price, S. L. Successful Computationally Directed Templating of Metastable Pharmaceutical Polymorphs. *Crystal Growth & Design* **2018**, *18* (9), 5322-5331.

(153) Cardew, P. T.; Davey, R. J. The kinetics of solvent-mediated phase transformations. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **1985**, *398* (1815), 415-428.

(154) Benavides, A. L.; Aragones, J. L.; Vega, C. Consensus on the solubility of NaCl in water from computer simulations using the chemical potential route. *The Journal of Chemical Physics* **2016**, *144* (12), 124504.

(155) Espinosa, J. R.; Young, J. M.; Jiang, H.; Gupta, D.; Vega, C.; Sanz, E.; Debenedetti, P. G.; Panagiotopoulos, A. Z. On the calculation of solubilities via direct coexistence simulations: Investigation of NaCl aqueous solutions and Lennard-Jones binary mixtures. *The Journal of Chemical Physics* **2016**, *145* (15), 154111.

(156) De Yoreo, J. J.; Zepeda-Ruiz, L. A.; Friddle, R. W.; Qiu, S. R.; Wasylenki, L. E.; Chernov, A. A.; Gilmer, G. H.; Dove, P. M. Rethinking Classical Crystal Growth Models through Molecular Scale Insights: Consequences of Kink-Limited Kinetics. *Crystal Growth & Design* **2009**, *9* (12), 5135-5144.

(157) Li, J. J.; Tilbury, C. J.; Joswiak, M. N.; Peters, B.; Doherty, M. F. Rate Expressions for Kink Attachment and Detachment During Crystal Growth. *Crystal Growth & Design* **2016**, *16* (6), 3313-3322.

(158) Stack, A. G. Molecular Dynamics Simulations of Solvation and Kink Site Formation at the {001} Barite−Water Interface. *The Journal of Physical Chemistry C* **2009**, *113* (6), 2104-2110.

(159) Joswiak, M. N.; Doherty, M. F.; Peters, B. Critical length of a one-dimensional nucleus. *The Journal of Chemical Physics* **2016**, *145* (21), 211916.

(160) De La Pierre, M.; Raiteri, P.; Stack, A. G.; Gale, J. D. Uncovering the Atomistic Mechanism for Calcite Step Growth. *Angewandte Chemie International Edition* **2017**, *56* (29), 8464-8467.

(161) Salvalaglio, M.; Perego, C.; Giberti, F.; Mazzotti, M.; Parrinello, M. Molecular-dynamics simulations of urea nucleation from aqueous solution. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112* (1), E6-E14.

(162) Kolafa, J. Solubility of NaCl in water and its melting point by molecular dynamics in the slab geometry and a new BK3-compatible force field. *The Journal of Chemical Physics* **2016**, *145* (20), 204509.

(163) Nielsen, A. E. *Kinetics of Precipitation*; Pergamon Press, 1964.

(164) Weeks, J. D.; Gilmer, G. H. Dynamics of Crystal Growth. *Advances in Chemical Physics* **1979**, 157-228.

(165) Chernov, A. A. Present-day understanding of crystal growth from aqueous solutions. *Progress in Crystal Growth and Characterization of Materials* **1993**, *26*, 121-151.

(166) Joswiak, M. N.; Doherty, M. F.; Peters, B. Ion dissolution mechanism and kinetics at kink sites on NaCl surfaces. *Proceedings of the National Academy of Sciences of the United States of America* **2018**, *115* (4), 656-661.

(167) Bjelobrk, Z.; Mendels, D.; Karmakar, T.; Parrinello, M.; Mazzotti, M. Solubility Prediction of Organic Molecules with Molecular Dynamics Simulations. *Crystal Growth & Design* **2021**, *21* (9), 5198-5205.

(168) Bjelobrk, Z.; Rajagopalan, A. K.; Mendels, D.; Karmakar, T.; Parrinello, M.; Mazzotti, M. Solubility of Organic Salts in Solvent–Antisolvent Mixtures: A Combined Experimental and Molecular Dynamics Simulations Approach. *Journal of Chemical Theory and Computation* **2022**, *18* (8), 4952-4959.

(169) Khanna, V.; Doherty, M. F.; Peters, B. Absolute chemical potentials for complex molecules in fluid phases: A centroid reference for predicting phase equilibria. *The Journal of Chemical Physics* **2020**, *153* (21), 214504.

(170) Khanna, V.; Doherty, M. F.; Peters, B. Predicting solubility and driving forces for crystallization using the absolute chemical potential route. *Molecular Physics* **2023**, *121* (2), e2155595.

(171) Widom, B. Some Topics in the Theory of Fluids. *The Journal of Chemical Physics* **1963**, *39* (11), 2808-2812.

(172) Siepmann, J. I.; Frenkel, D. Configurational bias Monte Carlo: a new sampling scheme for flexible chains. *Molecular Physics* **1992**, *75* (1), 59-70.

(173) Shi, W.; Maginn, E. J. Continuous Fractional Component Monte Carlo: An Adaptive Biasing Method for Open System Atomistic Simulations. *Journal of Chemical Theory and Computation* **2007**, *3* (4), 1451-1463.

(174) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids. *The Journal of Chemical Physics* **1971**, *54* (12), 5237-5247.

(175) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation* **2010**, *6* (5), 1509-1519.

(176) Mobley, D. L.; Guthrie, J. P. FreeSolv: a database of experimental and calculated hydration free energies, with input files. *Journal of Computer-Aided Molecular Design* **2014**, *28* (7), 711-720.

(177) Khanna, V.; Monroe, J. I.; Doherty, M. F.; Peters, B. Performing solvation free energy calculations in LAMMPS using the decoupling approach. *Journal of Computer-Aided Molecular Design* **2020**, *34* (6), 641-646.

(178) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *Journal of Physical Chemistry* **1987**, *91* (24), 6269-6271.

(179) Cappa, C. D.; Lovejoy, E. R.; Ravishankara, A. R. Determination of Evaporation Rates and Vapor Pressures of Very Low Volatility Compounds: A Study of the C4−C10 and C12 Dicarboxylic Acids. *The Journal of Physical Chemistry A* **2007**, *111* (16), 3099-3109.

(180) Bilde, M.; Svenningsson, B.; Mønster, J.; Rosenørn, T. Even−Odd Alternation of Evaporation Rates and Vapor Pressures of C3−C9 Dicarboxylic Acid Aerosols. *Environmental Science & Technology* **2003**, *37* (7), 1371-1378.

(181) Bilde, M.; Barsanti, K.; Booth, M.; Cappa, C. D.; Donahue, N. M.; Emanuelsson, E. U.; McFiggans, G.; Krieger, U. K.; Marcolli, C.; Topping, D.; et al. Saturation Vapor Pressures and Transition Enthalpies of Low-Volatility Organic Molecules of Atmospheric Relevance: From Dicarboxylic Acids to Complex Mixtures. *Chemical Reviews* **2015**, *115* (10), 4115-4156.

(182) Saleh, R.; Shihadeh, A.; Khlystov, A. Determination of evaporation coefficients of semi-volatile organic aerosols using an integrated volume—tandem differential mobility analysis (IV-TDMA) method. *Journal of Aerosol Science* **2009**, *40* (12), 1019-1029.

2191 (183) Yu, Q.; Black, S.; Wei, H. Solubility of Butanedioic Acid in Different Solvents at
2192 Temperatures between 283 K and 333 K. *Journal of Chemical & Engineering Data* **2009**, *54* (7),
2193 2123-2125.
2194 (184) Wick, C. D.; Siepmann, J. I.; Klotz, W. L.; Schure, M. R. Temperature effects on the
2195 retention of n-alkanes and arenes in helium–squalane gas–liquid chromatography: Experiment
2196 and molecular simulation. *Journal of Chromatography A* **2002**, *954* (1), 181-190.
2197 (185) Gillet, R.; Fierro, A.; Valenzuela, L. M.; Pérez-Correa, J. R. Using molecular dynamics
2198 simulations to predict the effect of temperature on aqueous solubility for aromatic compounds.
2199 *Fluid Phase Equilibria* **2018**, *472*, 85-93.
2200 (186) Ahmed, A.; Sandler, S. I. Temperature-Dependent Physicochemical Properties and
2201 Solvation Thermodynamics of Nitrotoluenes from Solvation Free Energies. *Journal of Chemical*
2202 *& Engineering Data* **2015**, *60* (1), 16-27.
2203 (187) Grant, D. J. W.; Mehdizadeh, M.; Chow, A. H. L.; Fairbrother, J. E. Non-linear van't Hoff
2204 solubility-temperature plots and their pharmaceutical interpretation. *International Journal of*
2205 *Pharmaceutics* **1984**, *18* (1), 25-38.
2206 (188) Fowles, D. J.; Palmer, D. S.; Guo, R.; Price, S. L.; Mitchell, J. B. O. Toward Physics-
2207 Based Solubility Computation for Pharmaceuticals to Rival Informatics. *Journal of Chemical*
2208 *Theory and Computation* **2021**, *17* (6), 3700-3709.
2209 (189) Rees, D. C.; Wolfe, G. M. Macromolecular solvation energies derived from small-
2210 molecule crystal morphology. *Protein Science* **1993**, *2* (11), 1882-1889.
2211 (190) da Silva, M. A. V.; Monte, M. J. S.; Ribeiro, J. R. Thermodynamic study on the
2212 sublimation of succinic acid and of methyl- and dimethyl-substituted succinic and glutaric acids.
2213 *Journal of Chemical Thermodynamics* **2001**, *33* (1), 23-31.
2214 (191) Homeyer, N.; Gohlke, H. Free Energy Calculations by the Molecular Mechanics
2215 Poisson−Boltzmann Surface Area Method. *Molecular Informatics* **2012**, *31* (2), 114-122.
2216 (192) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. Ab initio study of solvated molecules: a new
2217 implementation of the polarizable continuum model. *Chemical Physics Letters* **1996**, *255* (4),
2218 327-335.
2219 (193) Mullins, E.; Oldland, R.; Liu, Y. A.; Wang, S.; Sandler, S. I.; Chen, C.-C.; Zwolak, M.;
2220 Seavey, K. C. Sigma-Profile Database for Using COSMO-Based Thermodynamic Methods.
2221 *Industrial & Engineering Chemistry Research* **2006**, *45* (12), 4389-4415.
2222 (194) Abramov, Y. A.; Sun, G. X.; Zeng, Q.; Yang, M. J. Guiding Lead Optimization for
2223 Solubility Improvement with Physics-Based Modeling. *Molecular Pharmaceutics* **2020**, *17* (2),
2224 666-673.
2225 (195) Median (50th percentile) particle diameter for a volumetric particle size distribution
2226 assuming spherical particles. More generally Xv50 where X is a characteristic length for a
2227 volumetric particle size distribution.
2228 (196) Shayesteh Zadeh, A.; Peters, B. Conformational Interconversion Kinetics, Boundary Layer
2229 Transport, and Crystal Growth Impedance. *Crystal Growth & Design* **2022**, *22* (7), 4298-4304.
2230 (197) Vekilov, P. G. Nucleation. *Crystal Growth & Design* **2010**, *10* (12), 5007-5019.
2231 (198) Burton, W. K.; Cabrera, N.; Frank, F. C. The growth of crystals and the equilibrium
2232 structure of their surfaces. *Philosophical Transactions of the Royal Society of London Series A-*
2233 *Mathematical Physical and Engineering Sciences* **1951**, *243* (866), 299-358.
2234 (199) Chernov, A. A. THE SPIRAL GROWTH OF CRYSTALS. *Soviet Physics Uspekhi* **1961**, *4*
2235 (1), 116-148.

2236 (200) Ohara, M.; Reid, R. C. *Modelling Crystal Growth Rates from Solution*; Prentice-Hall,
2237 1973.
2238 (201) Chernov, A. A. *Modern Crystallography III, Crystal Growth*; Springer-Verlag, 1984.
2239 (202) Lovette, M. A.; Doherty, M. F. Reinterpreting edge energies calculated from crystal
2240 growth experiments. *Journal of Crystal Growth* **2011**, *327* (1), 117-126.
2241 (203) Tilbury, C. J.; Doherty, M. F. Modeling layered crystal growth at increasing
2242 supersaturation by connecting growth regimes. *AIChE Journal* **2017**, *63* (4), 1338-1352.
2243 (204) Lovette, M. A.; Browning, A. R.; Griffin, D. W.; Sizemore, J. P.; Snyder, R. C.; Doherty,
2244 M. F. Supporting Information in: Crystal Shape Engineering. *Industrial & Engineering*
2245 *Chemistry Research* **2008**, *47* (24), 9812-9833.
2246 (205) Snyder, R. C.; Doherty, M. F. Predicting crystal growth by spiral motion. *Proceedings of*
2247 *the Royal Society A: Mathematical, Physical and Engineering Sciences* **2009**, *465* (2104), 1145-
2248 1171.
2249 (206) Teng, H. H.; Dove, P. M.; Orme, C. A.; De Yoreo, J. J. Thermodynamics of Calcite
2250 Growth: Baseline for Understanding Biomineral Formation. *Science* **1998**, *282* (5389), 724-727.
2251 (207) Voronkov, V. V. Dislocation mechanism of growth with a low kink density. *Soviet Physics*
2252 *Crystallography* **1973**, *18* (1), 19-23.
2253 (208) Frenkel, J. ON THE SURFACE MOTION OF PARTICLES IN CRYSTALS AND THE
2254 NATURAL ROUGHNESS OF CRYSTALLINE FACES. *Journal of Physics, Academy of*
2255 *Sciences of the USSR* **1945**, *9* (5), 392-398.
2256 (209) Israelachvili, J. N. Chapter 17. In *Intermolecular and Surface Forces*, Academic Press,
2257 2011.
2258 (210) Winn, D.; Doherty, M. F. A new technique for predicting the shape of solution-grown
2259 organic crystals. *AIChE Journal* **1998**, *44* (11), 2501-2514.
2260 (211) Winn, D.; Doherty, M. F. Predicting the shape of organic crystals grown from polar
2261 solvents. *Chemical Engineering Science* **2002**, *57* (10), 1805-1813.
2262 (212) Tilbury, C. J.; Green, D. A.; Marshall, W. J.; Doherty, M. F. Predicting the Effect of
2263 Solvent on the Crystal Habit of Small Organic Molecules. *Crystal Growth & Design* **2016**, *16*
2264 (5), 2590-2604.
2265 (213) van der Eerden, J. P. Crystal growth mechanisms. In *Handbook of Crystal Growth;*
2266 *Fundamentals, Transport and Stability*, Hurle, D. T. J. Ed.; North-Holland, 1993.
2267 (214) Tilbury, C. J.; Joswiak, M. N.; Peters, B.; Doherty, M. F. Modeling Step Velocities and
2268 Edge Surface Structures during Growth of Non-Centrosymmetric Crystals. *Crystal Growth &*
2269 *Design* **2017**, *17* (4), 2066-2080.
2270 (215) Sun, Y. Y.; Tilbury, C. J.; Reutzel-Edens, S. M.; Bhardwaj, R. M.; Li, J. J.; Doherty, M. F.
2271 Modeling Olanzapine Solution Growth Morphologies. *Crystal Growth & Design* **2018**, *18* (2),
2272 905-911.
2273 (216) Sours, R. E.; Zellelow, A. Z.; Swift, J. A. An in Situ Atomic Force Microscopy Study of
2274 Uric Acid Crystal Growth. *The Journal of Physical Chemistry B* **2005**, *109* (20), 9989-9995.
2275 (217) Vekilov, P. G. What Determines the Rate of Growth of Crystals from Solution? *Crystal*
2276 *Growth & Design* **2007**, *7* (12), 2796-2810.
2277 (218) Rimer, J. D.; An, Z.; Zhu, Z.; Lee, M. H.; Goldfarb, D. S.; Wesson, J. A.; Ward, M. D.
2278 Crystal Growth Inhibitors for the Prevention of l-Cystine Kidney Stones Through Molecular
2279 Design. *Science* **2010**, *330* (6002), 337-341.
2280 (219) Hartman, P.; Perdok, W. G. On the Relations Between Structure and Morphology of
2281 Crystals. I. *Acta Crystallographica* **1955**, *8* (1), 49-52.

2282 (220) Hartman, P.; Perdok, W. G. On the Relations Between Structure and Morphology of
2283 Crystals. II. *Acta Crystallographica* **1955**, *8* (9), 521-524.
2284 (221) Hartman, P.; Perdok, W. G. On the Relations Between Structure and Morphology of
2285 Crystals. III. *Acta Crystallographica* **1955**, *8* (9), 525-529.
2286 (222) Gibbs, J. W. *The Scientific Papers of J. Willard Gibbs: Volume One - Thermodynamics*;
2287 Dover Publications, 1961.
2288 (223) Frank, F. C. On the kinematic theory of crystal growth and dissolution processes. In
2289 *Growth and Perfection of Crystals*, Doremus, R. H., Roberts, B. W., Turnbull, D. Eds.; John
2290 Wiley and Sons, 1958; pp 411-419.
2291 (224) Chernov, A. A. The kinetics of the growth forms of crystals. *Soviet Physics
2292 Crystallography* **1963**, *7*, 728-730.
2293 (225) Zhang, Y.; Sizemore, J. P.; Doherty, M. F. Shape evolution of 3-dimensional faceted
2294 crystals. *AIChE Journal* **2006**, *52* (5), 1906-1915.
2295 (226) Snyder, R. C.; Doherty, M. F. Faceted crystal shape evolution during dissolution or
2296 growth. *AIChE Journal* **2007**, *53* (5), 1337-1348.
2297 (227) Li, J.; Tilbury, C. J.; Kim, S. H.; Doherty, M. F. A design aid for crystal growth
2298 engineering. *Progress in Materials Science* **2016**, *82*, 1-38.
2299 (228) Lovette, M. A.; Doherty, M. F. Predictive Modeling of Supersaturation-Dependent Crystal
2300 Shapes. *Crystal Growth & Design* **2012**, *12* (2), 656-669.
2301 (229) Gavezzotti, A. Efficient computer modeling of organic materials. The atom-atom,
2302 Coulomb-London-Pauli (AA-CLP) model for intermolecular electrostatic-polarization,
2303 dispersion and repulsion energies. *New Journal of Chemistry* **2011**, *35* (7), 1360-1368.
2304 (230) Van Oss, C. J.; Chaudhury, M. K.; Good, R. J. Interfacial Lifshitz-van der Waals and polar
2305 interactions in macroscopic systems. *Chemical Reviews* **1988**, *88* (6), 927-941.
2306 (231) Davey, R. J.; Black, S. N.; Logan, D.; Maginn, S. J.; Fairbrother, J. E.; Grant, D. J. W.
2307 Structural and kinetic features of crystal growth inhibition: adipic acid growing in the presence
2308 of n-alkanoic acids. *Journal of the Chemical Society, Faraday Transactions* **1992**, *88* (23), 3461-
2309 3466.
2310 (232) Wawrzycka-Gorczyca, I.; Koziol, A. E.; Glice, M.; Cybulski, J. Polymorphic form II of 2-
2311 methyl-4-(4-methyl-1-piperazinyl)-10H-thieno[2,3-b][1,5]benzodiazepine. *Acta
2312 Crystallographica Section E - Structure Reports Online* **2004**, *60*, o66-o68.
2313 (233) Wawrzycka-Gorczyca, I.; Mazur, L.; Koziol, A. E. 2-methyl-4-(4-methyl-1-piperazinyl)-
2314 10H-thieno[2,3-b][1,5]benzodiazepine methanol solvate. *Acta Crystallographica Section E -
2315 Structure Reports Online* **2004**, *60*, O69-O71.
2316 (234) Wawrzycka-Gorczyca, I.; Borowski, P.; Osypiuk-Tomasik, J.; Mazur, L.; Koziol, A. E.
2317 Crystal structure of olanzapine and its solvates. Part 3. Two and three-component solvates with
2318 water, ethanol, butan-2-ol and dichloromethane. *Journal of Molecular Structure* **2007**, *830* (1-3),
2319 188-197.
2320 (235) Verma, L.; Warzecha, M.; Chakrabarti, R.; Hadjiev, V. G.; Palmer, J. C.; Vekilov, P. G.
2321 How to Identify the Crystal Growth Unit. *Israel Journal of Chemistry* **2021**, *61* (11-12), 818-827.
2322 (236) Cuppen, H. M.; Meekes, H.; van Veenendaal, E.; van Enckevort, W. J. P.; Bennema, P.;
2323 Reedijk, M. F.; Arsic, J.; Vlieg, E. Kink density and propagation velocity of the [010] step on the
2324 Kossel (100) surface. *Surface Science* **2002**, *506* (3), 183-195.
2325 (237) Padwal, N. A.; Doherty, M. F. Simple Accurate Nonequilibrium Step Velocity Model for
2326 Crystal Growth of Symmetric Organic Molecules. *Crystal Growth & Design* **2022**, *22* (6), 3656-
2327 3661.

2328 (238) Zhang, J.; Nancollas, G. H. Kink Density and Rate of Step Movement during Growth and
2329 Dissolution of anABCrystal in a Nonstoichiometric Solution. *Journal of Colloid and Interface*
2330 *Science* **1998**, *200* (1), 131-145.
2331 (239) Kuvadia, Z. B.; Doherty, M. F. Spiral Growth Model for Faceted Crystals of Non-
2332 Centrosymmetric Organic Molecules Grown from Solution. *Crystal Growth & Design* **2011**, *11*
2333 (7), 2780-2802.
2334 (240) Cuppen, H. M.; Meekes, H.; van Enckevort, W. J. P.; Vlieg, E. Kink incorporation and
2335 step propagation in a non-Kossel model. *Surface Science* **2004**, *571* (1), 41-62.
2336 (241) Joswiak, M. N.; Peters, B.; Doherty, M. F. Crystal step edges with alternating rows of
2337 growth units: 1D nucleation and step velocity. *Journal of Crystal Growth* **2023**, *604*, 127042.
2338 (242) Shim, H.-M.; Koo, K.-K. Crystal Morphology Prediction of Hexahydro-1,3,5-trinitro-
2339 1,3,5-triazine by the Spiral Growth Model. *Crystal Growth & Design* **2014**, *14* (4), 1802-1810.
2340 (243) Shim, H.-M.; Kim, H.-S.; Koo, K.-K. Molecular Modeling on Supersaturation-Dependent
2341 Growth Habit of 1,1-Diamino-2,2-dinitroethylene. *Crystal Growth & Design* **2015**, *15* (4), 1833-
2342 1842.
2343 (244) Shim, H.-M.; Koo, K.-K. Molecular Approach to the Effect of Interfacial Energy on
2344 Growth Habit of ε-HNIW. *Crystal Growth & Design* **2016**, *16* (11), 6506-6513.
2345 (245) Zhao, Y. S.; Tilbury, C. J.; Landis, S.; Sun, Y. Y.; Li, J. J.; Zhu, P.; Doherty, M. F. A New
2346 Software Framework for Implementing Crystal Growth Models to Materials of Any
2347 Crystallographic Complexity. *Crystal Growth & Design* **2020**, *20* (5), 2885-2892.
2348 (246) Hill, A. R.; Cubillas, P.; Gebbie-Rayet, J. T.; Trueman, M.; de Bruyn, N.; al Harthi, Z.;
2349 Pooley, R. J. S.; Attfield, M. P.; Blatov, V. A.; Proserpio, D. M.; et al. CrystalGrower: a generic
2350 computer program for Monte Carlo modelling of crystal growth. *Chemical Science* **2021**, *12* (3),
2351 1126-1146.
2352 (247) Stack, A. G.; Raiteri, P.; Gale, J. D. Accurate Rates of the Complex Mechanisms for
2353 Growth and Dissolution of Minerals Using a Combination of Rare-Event Theories. *Journal of*
2354 *the American Chemical Society* **2012**, *134* (1), 11-14.
2355 (248) Joswiak, M. N.; Peters, B.; Doherty, M. F. In Silico Crystal Growth Rate Prediction for
2356 NaCl from Aqueous Solution. *Crystal Growth & Design* **2018**, *18* (10), 6302-6306.
2357 (249) Dighe, A. V.; Singh, M. R. Solvent fluctuations in the solvation shell determine the
2358 activation barrier for crystal growth rates. *Proceedings of the National Academy of Sciences*
2359 **2019**, *116* (48), 23954-23959.
2360 (250) Kuvadia, Z. B.; Doherty, M. F. Reformulating multidimensional population balances for
2361 predicting crystal size and shape. *AIChE Journal* **2013**, *59* (9), 3468-3474.
2362 (251) Marinova, V.; Wood, G. P. F.; Marziano, I.; Salvalaglio, M. Dynamics and
2363 Thermodynamics of Ibuprofen Conformational Isomerism at the Crystal/Solution Interface.
2364 *Journal of Chemical Theory and Computation* **2018**, *14* (12), 6484-6494.
2365 (252) Jiang, H.; Haji-Akbari, A.; Debenedetti, P. G.; Panagiotopoulos, A. Z. Forward flux
2366 sampling calculation of homogeneous nucleation rates from aqueous NaCl solutions. *The*
2367 *Journal of Chemical Physics* **2018**, *148* (4), 044505.
2368 (253) Lanaro, G.; Patey, G. N. Birth of NaCl Crystals: Insights from Molecular Simulations. *The*
2369 *Journal of Physical Chemistry B* **2016**, *120* (34), 9076-9087.
2370 (254) Salvalaglio, M.; Mazzotti, M.; Parrinello, M. Urea homogeneous nucleation mechanism is
2371 solvent dependent. *Faraday Discussions* **2015**, *179* (0), 291-307.

(255) Liu, C. X.; Cao, F. J.; Kulkarni, S. A.; Wood, G. P. F.; Santiso, E. E. Understanding Polymorph Selection of Sulfamerazine in Solution. *Crystal Growth & Design* **2019**, *19* (12), 6925-6934.

(256) Anwar, J.; Zahn, D. Uncovering Molecular Processes in Crystal Nucleation and Growth by Using Molecular Simulation. *Angewandte Chemie-International Edition* **2011**, *50* (9), 1996-2013.

(257) Poon, G. G.; Peters, B. Accelerated Nucleation Due to Trace Additives: A Fluctuating Coverage Model. *The Journal of Physical Chemistry B* **2016**, *120* (8), 1679-1684.

(258) Poon, G. G.; Lemke, T.; Peter, C.; Molinero, V.; Peters, B. Soluble Oligomeric Nucleants: Simulations of Chain Length, Binding Strength, and Volume Fraction Effects. *The Journal of Physical Chemistry Letters* **2017**, *8* (23), 5815-5820.

(259) Kumar, A.; Molinero, V. Two-Step to One-Step Nucleation of a Zeolite through a Metastable Gyroid Mesophase. *The Journal of Physical Chemistry Letters* **2018**, *9* (19), 5692-5697.

(260) Palafox-Hernandez, J. P.; Laird, B. B. Orientation dependence of heterogeneous nucleation at the Cu–Pb solid-liquid interface. *The Journal of Chemical Physics* **2016**, *145* (21), 211914.

(261) Peters, B. *Reaction Rate Theory and Rare Events Simulations*; Elsevier, 2017. DOI: https://doi.org/10.1016/B978-0-44-456349-1.00025-8.

(262) Hall, S. W.; Díaz Leines, G.; Sarupria, S.; Rogal, J. Practical guide to replica exchange transition interface sampling and forward flux sampling. *The Journal of Chemical Physics* **2022**, *156* (20), 200901.

(263) ten Wolde, P. R.; Frenkel, D. Computer simulation study of gas–liquid nucleation in a Lennard-Jones system. *The Journal of Chemical Physics* **1998**, *109* (22), 9901-9918.

(264) ten Wolde, P. R.; Ruiz-Montero, M. J.; Frenkel, D. Numerical calculation of the rate of crystal nucleation in a Lennard-Jones system at moderate undercooling. *The Journal of Chemical Physics* **1996**, *104* (24), 9932-9947.

(265) Auer, S.; Frenkel, D. QUANTITATIVE PREDICTION OF CRYSTAL-NUCLEATION RATES FOR SPHERICAL COLLOIDS: A Computational Approach. *Annual Review of Physical Chemistry* **2004**, *55* (1), 333-361.

(266) Salvalaglio, M.; Tiwary, P.; Maggioni, G. M.; Mazzotti, M.; Parrinello, M. Overcoming time scale and finite size limitations to compute nucleation rates from small scale well tempered metadynamics simulations. *The Journal of Chemical Physics* **2016**, *145* (21), 211925.

(267) Agarwal, V.; Peters, B. Solute Precipitate Nucleation: A Review of Theory and Simulation Advances. In *Advances in Chemical Physics: Volume 155*, Advances in Chemical Physics, 2014; pp 97-160.

(268) Fan, Z.; Men, H. An Overview on Atomistic Mechanisms of Heterogeneous Nucleation. In *Metals*, 2022; Vol. 12, p 1547.

(269) Davey, R. J.; Schroeder, S. L.; ter Horst, J. H. Nucleation of Organic CrystalsA Molecular Perspective. *Angewandte Chemie-International Edition* **2013**, *52* (8), 2166-2179.

(270) Cheng, B.; Tribello, G. A.; Ceriotti, M. Solid-liquid interfacial free energy out of equilibrium. *Physical Review B* **2015**, *92* (18), 180102.

(271) Karthika, S.; Radhakrishnan, T. K.; Kalaichelvi, P. A Review of Classical and Nonclassical Nucleation Theories. *Crystal Growth & Design* **2016**, *16* (11), 6663-6681.

(272) *Nucleation in Condensed Matter: Applications in Materials and Biology*; Pergamon, 2010.

(273) Kashchiev, D. *Nucleation*; Butterworth-Heinemann, 2000.

(274) Zimmermann, N. E. R.; Vorselaars, B.; Espinosa, J. R.; Quigley, D.; Smith, W. R.; Sanz, E.; Vega, C.; Peters, B. NaCl nucleation from brine in seeded simulations: Sources of uncertainty in rate estimates. *The Journal of Chemical Physics* **2018**, *148* (22), 222838.

(275) Knott, B. C.; Molinero, V.; Doherty, M. F.; Peters, B. Homogeneous Nucleation of Methane Hydrates: Unrealistic under Realistic Conditions. *Journal of the American Chemical Society* **2012**, *134* (48), 19544-19547.

(276) Duff, N.; Peters, B. Polymorph specific RMSD local order parameters for molecular crystals and nuclei: a-, b-, and c- glycine. *Journal of Chemical Physics* **2011**, *135* (13), 134101.

(277) Knott, B. C.; Duff, N.; Doherty, M. F.; Peters, B. Estimating diffusivity along a reaction coordinate in the high friction limit: Insights on pulse times in laser-induced nucleation. *The Journal of Chemical Physics* **2009**, *131* (22), 224112.

(278) *Nucleation Theory and Applications*; WILEY-VCH Verlag GmbH & Co.KGaA,, 2005.

(279) Kalikmanov, V. I. *Nucleation Theory*; Springer, 2013.

(280) Garside, J.; Davey, R. J. INVITED REVIEW SECONDARY CONTACT NUCLEATION: KINETICS, GROWTH AND SCALE-UP. *Chemical Engineering Communications* **1980**, *4* (4-5), 393-424.

(281) Hounslow, M. J. Nucleation, growth, and aggregation rates from steady-state experimental data. *AIChE Journal* **1990**, *36* (11), 1748-1752.

(282) Moučka, F.; Nezbeda, I.; Smith, W. R. Chemical Potentials, Activity Coefficients, and Solubility in Aqueous NaCl Solutions: Prediction by Polarizable Force Fields. *Journal of Chemical Theory and Computation* **2015**, *11* (4), 1756-1764.

(283) Nezbeda, I.; Moučka, F.; Smith, W. R. Recent progress in molecular simulation of aqueous electrolytes: force fields, chemical potentials and solubility. *Molecular Physics* **2016**, *114* (11), 1665-1690.

(284) Mester, Z.; Panagiotopoulos, A. Z. Temperature-dependent solubilities and mean ionic activity coefficients of alkali halides in water from molecular dynamics simulations. *The Journal of Chemical Physics* **2015**, *143* (4), 044505.

(285) Joswiak, M. N.; Duff, N.; Doherty, M. F.; Peters, B. Size-Dependent Surface Free Energy and Tolman-Corrected Droplet Nucleation of TIP4P/2005 Water. *The Journal of Physical Chemistry Letters* **2013**, *4* (24), 4267-4272.

(286) Heerman, D. W. Classical nucleation theory with a Tolman correction *J. Stat. Phys.* **1982**, *29*, 631-640.

(287) Bertolazzo, A. A.; Naullage, P. M.; Peters, B.; Molinero, V. The Clathrate–Water Interface Is Oleophilic. *The Journal of Physical Chemistry Letters* **2018**, *9* (12), 3224-3231.

(288) Punnathanam, S.; Monson, P. A. Crystal nucleation in binary hard sphere mixtures: A Monte Carlo simulation study. *The Journal of Chemical Physics* **2006**, *125* (2), 024508.

(289) Lifanov, Y.; Vorselaars, B.; Quigley, D. Nucleation barrier reconstruction via the seeding method in a lattice model with competing nucleation pathways. *The Journal of Chemical Physics* **2016**, *145* (21), 211912.

(290) Sanz, E.; Vega, C.; Espinosa, J. R.; Caballero-Bernal, R.; Abascal, J. L. F.; Valeriani, C. Homogeneous Ice Nucleation at Moderate Supercooling from Molecular Simulation. *Journal of the American Chemical Society* **2013**, *135* (40), 15008-15017.

(291) Liu, C. X.; Wood, G. P. F.; Santiso, E. E. Modelling nucleation from solution with the string method in the osmotic ensemble. *Molecular Physics* **2018**, *116* (21-22), 2998-3007.

(292) ten Wolde, P. R.; Frenkel, D. Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations. *Science* **1997**, *277* (5334), 1975-1978.

2463 (293) Vekilov, P. G. Two-step mechanism for the nucleation of crystals from solution. *Journal of*
2464 *Crystal Growth* **2005**, *275* (1), 65-76.
2465 (294) Bosetti, L.; Mazzotti, M. Population Balance Modeling of Growth and Secondary
2466 Nucleation by Attrition and Ripening. *Crystal Growth & Design* **2020**, *20* (1), 307-319.
2467 (295) Ahn, B.; Bosetti, L.; Mazzotti, M. Secondary Nucleation by Interparticle Energies. III.
2468 Nucleation Rate Model. *Crystal Growth & Design* **2022**, *22* (6), 3625-3636.
2469 (296) Xu, S.; Hou, Z.; Chuai, X.; Wang, Y. Overview of Secondary Nucleation: From
2470 Fundamentals to Application. *Industrial & Engineering Chemistry Research* **2020**, *59* (41),
2471 18335-18356.
2472 (297) *Industrial Crystallization*; Springer, 1976.
2473 (298) Hoffmann, J.; Flannigan, J.; Cashmore, A.; Briuglia, Maria L.; Steendam, R. R. E.; Gerard,
2474 C. J. J.; Haw, M. D.; Sefcik, J.; ter Horst, J. H. The unexpected dominance of secondary over
2475 primary nucleation. *Faraday Discussions* **2022**, *235* (0), 109-131.
2476 (299) Bal, V.; Peters, B. Crystallization with Sinusoidal Modulation of Stirrer Speed: Frequency
2477 Response Analysis and Secondary Nucleation Kinetics. *Crystal Growth & Design* **2021**, *21* (1),
2478 235-242.
2479 (300) Kiely, E.; Zwane, R.; Fox, R.; Reilly, A. M.; Guerin, S. Density functional theory
2480 predictions of the mechanical properties of crystalline materials. *Crystengcomm* **2021**, *23* (34),
2481 5697-5710.
2482 (301) *Crystallization Technology Handbook*; Marcel Dekker, Inc., 2001.
2483 (302) Turnbull, D. Phase Changes. In *Solid State Physics*, Seitz, F., Turnbull, D. Eds.; Vol. 3;
2484 Academic Press, 1956; pp 225-306.
2485 (303) Cabriolu, R.; Li, T. Ice nucleation on carbon surface supports the classical theory for
2486 heterogeneous nucleation. *Physical Review E* **2015**, *91* (5), 052402.
2487 (304) Auer, S.; Frenkel, D. Line Tension Controls Wall-Induced Crystal Nucleation in Hard-
2488 Sphere Colloids. *Physical Review Letters* **2003**, *91* (1), 015703.
2489 (305) Navascués, G.; Tarazona, P. Line tension effects in heterogeneous nucleation theory. *The*
2490 *Journal of Chemical Physics* **1981**, *75* (5), 2441-2446.
2491 (306) Mithen, J. P.; Sear, R. P. Computer simulation of epitaxial nucleation of a crystal on a
2492 crystalline surface. *The Journal of Chemical Physics* **2014**, *140* (8), 084504.
2493 (307) Turnbull, D.; Vonnegut, B. Nucleation Catalysis. *Industrial & Engineering Chemistry*
2494 **1952**, *44* (6), 1292-1298.
2495 (308) Fletcher, N. H. Size Effect in Heterogeneous Nucleation. *The Journal of Chemical Physics*
2496 **1958**, *29* (3), 572-576.
2497 (309) Lupi, L.; Hudait, A.; Molinero, V. Heterogeneous Nucleation of Ice on Carbon Surfaces.
2498 *Journal of the American Chemical Society* **2014**, *136* (8), 3156-3164.
2499 (310) Fitzner, M.; Sosso, G. C.; Cox, S. J.; Michaelides, A. The Many Faces of Heterogeneous
2500 Ice Nucleation: Interplay Between Surface Morphology and Hydrophobicity. *Journal of the*
2501 *American Chemical Society* **2015**, *137* (42), 13658-13669.
2502 (311) Lupi, L.; Molinero, V. Does Hydrophilicity of Carbon Particles Improve Their Ice
2503 Nucleation Ability? *The Journal of Physical Chemistry A* **2014**, *118* (35), 7330-7337.
2504 (312) Sosso, G. C.; Li, T.; Donadio, D.; Tribello, G. A.; Michaelides, A. Microscopic
2505 Mechanism and Kinetics of Ice Formation at Complex Interfaces: Zooming in on Kaolinite. *The*
2506 *Journal of Physical Chemistry Letters* **2016**, *7* (13), 2350-2355.

2507   (313) Lochhead, M. J.; Letellier, S. R.; Vogel, V. Assessing the Role of Interfacial Electrostatics
2508   in Oriented Mineral Nucleation at Charged Organic Monolayers. *The Journal of Physical*
2509   *Chemistry B* **1997**, *101* (50), 10821-10827.
2510   (314) DeFever, R. S.; Sarupria, S. Surface chemistry effects on heterogeneous clathrate hydrate
2511   nucleation: A molecular dynamics study. *The Journal of Chemical Thermodynamics* **2018**, *117*,
2512   205-213.
2513   (315) Glatz, B.; Sarupria, S. Heterogeneous Ice Nucleation: Interplay of Surface Properties and
2514   Their Impact on Water Orientations. *Langmuir* **2018**, *34* (3), 1190-1198.
2515   (316) Cantor, B. Heterogeneous nucleation and adsorption. *Philosophical Transactions of the*
2516   *Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **2003**, *361*
2517   (1804), 409-417.
2518   (317) Bonafede, S. J.; Ward, M. D. Selective Nucleation and Growth of An Organic Polymorph
2519   by Ledge-Directed Epitaxy on A Molecular-Crystal Substrate. *Journal of the American Chemical*
2520   *Society* **1995**, *117* (30), 7853-7861.
2521   (318) Arlin, J. B.; Price, L. S.; Price, S. L.; Florence, A. J. A strategy for producing predicted
2522   polymorphs: catemeric carbamazepine form V. *Chemical Communications* **2011**, *47* (25), 7074-
2523   7076.
2524   (319) Chadwick, K.; Chen, J.; Myerson, A. S.; Trout, B. L. Toward the Rational Design of
2525   Crystalline Surfaces for Heteroepitaxy: Role of Molecular Functionality. *Crystal Growth &*
2526   *Design* **2012**, *12* (3), 1159-1166.
2527   (320) Parambil, J. V.; Poornachary, S. K.; Heng, J. Y. Y.; Tan, R. B. H. Template-induced
2528   nucleation for controlling crystal polymorphism: from molecular mechanisms to applications in
2529   pharmaceutical processing. *CrystEngComm* **2019**, *21* (28), 4122-4135.
2530   (321) Mitchell, C. A.; Yu, L.; Ward, M. D. Selective nucleation and discovery of organic
2531   polymorphs through epitaxy with single crystal substrates. *Journal of the American Chemical*
2532   *Society* **2001**, *123* (44), 10830-10839.
2533   (322) Kim, K.; Lee, I. S.; Centrone, A.; Hatton, T. A.; Myerson, A. S. Formation of Nanosized
2534   Organic Molecular Crystals on Engineered Surfaces. *Journal of the American Chemical Society*
2535   **2009**, *131* (51), 18212-+.
2536   (323) Ward, M. D. Soft Crystals in Flatland: Unraveling Epitaxial Growth. *ACS Nano* **2016**, *10*
2537   (7), 6424-6428.
2538   (324) Tan, L.; Davis, R. M.; Myerson, A. S.; Trout, B. L. Control of Heterogeneous Nucleation
2539   via Rationally Designed Biocompatible Polymer Surfaces with Nanoscale Features. *Crystal*
2540   *Growth & Design* **2015**, *15* (5), 2176-2186.
2541   (325) Stojakovic, J.; Baftizadeh, F.; Bellucci, M. A.; Myerson, A. S.; Trout, B. L. Angle-
2542   Directed Nucleation of Paracetamol on Biocompatible Nanoimprinted Polymers. *Crystal Growth*
2543   *& Design* **2017**, *17* (6), 2955-2963.
2544   (326) Lancaster, R. W.; Harris, L. D.; Pearson, D. Fifty-year old samples of progesterone
2545   demonstrate the complex role of synthetic impurities in stabilizing a metastable polymorph.
2546   *CrystEngComm* **2011**, *13* (6), 1775-1777.
2547   (327) Meldrum, F. C.; O'Shaughnessy, C. Crystallization in Confinement. *Advanced Materials*
2548   **2020**, *32* (31), 2001068.
2549   (328) Page, A. J.; Sear, R. P. Heterogeneous Nucleation in and out of Pores. *Physical Review*
2550   *Letters* **2006**, *97* (6), 065701.

2551 (329) Holden, M. A.; Whale, T. F.; Tarn, M. D.; O'Sullivan, D.; Walshaw, R. D.; Murray, B. J.;
2552 Meldrum, F. C.; Christenson, H. K. High-speed imaging of ice nucleation in water proves the
2553 existence of active sites. *Science Advances* **2019**, *5* (2), eaav4316.
2554 (330) Navrotsky, A. Energetic clues to pathways to biomineralization: Precursors, clusters, and
2555 nanoparticles. *Proceedings of the National Academy of Sciences* **2004**, *101* (33), 12096-12101.
2556 (331) Gimondi, I.; Salvalaglio, M. CO2 packing polymorphism under confinement in cylindrical
2557 nanopores. *Molecular Systems Design & Engineering* **2018**, *3* (1), 243-252.
2558 (332) Zhao, W.-H.; Wang, L.; Bai, J.; Yuan, L.-F.; Yang, J.; Zeng, X. C. Highly Confined
2559 Water: Two-Dimensional Ice, Amorphous Ice, and Clathrate Hydrates. *Accounts of Chemical*
2560 *Research* **2014**, *47* (8), 2505-2513.
2561 (333) Chayen, N. E.; Saridakis, E.; Sear, R. P. Experiment and theory for heterogeneous
2562 nucleation of protein crystals in a porous medium. *Proceedings of the National Academy of*
2563 *Sciences* **2006**, *103* (3), 597-601.
2564 (334) Grzesiak, A. L.; Uribe, F. J.; Ockwig, N. W.; Yaghi, O. M.; Matzger, A. J. Polymer-
2565 induced heteronucleation for the discovery of new extended solids. *Angewandte Chemie-*
2566 *International Edition* **2006**, *45* (16), 2553-2556.
2567 (335) Lopez-Mejias, V.; Knight, J. L.; Brooks, C. L.; Matzger, A. J. On the Mechanism of
2568 Crystalline Polymorph Selection by Polymer Heteronuclei. *Langmuir* **2011**, *27* (12), 7575-7579.
2569 (336) Pfund, L. Y.; Matzger, A. J. Towards Exhaustive and Automated High-Throughput
2570 Screening for Crystalline Polymorphs. *ACS Combinatorial Science* **2014**, *16* (7), 309-313.
2571 (337) Price, C. P.; Grzesiak, A. L.; Matzger, A. J. Crystalline polymorph selection and discovery
2572 with polymer heteronuclei. *Journal of the American Chemical Society* **2005**, *127* (15), 5512-
2573 5517.
2574 (338) Diao, Y.; Myerson, A. S.; Hatton, T.; Trout, B. L. Surface Design for Controlled
2575 Crystallization: The Role of Surface Chemistry and Nanoscale Pores in Heterogeneous
2576 Nucleation. *Langmuir* **2011**, *27* (9), 5324-5334.
2577 (339) Sear, R. P. Non-self-averaging nucleation rate due to quenched disorder. *Journal of*
2578 *Physics: Condensed Matter* **2011**, *24* (5), 052205.
2579 (340) Finney, A. R.; Salvalaglio, M. Multiple pathways in NaCl homogeneous crystal nucleation.
2580 *Faraday Discussions* **2022**, *235* (0), 56-80.
2581 (341) Tang, S. K.; Davey, R. J.; Sacchi, P.; Cruz-Cabeza, A. J. Can molecular flexibility control
2582 crystallization? The case of para substituted benzoic acids. *Chemical Science* **2021**, *12* (3), 993-
2583 1000.
2584 (342) Aoun, M.; Plasari, E.; David, R.; Villermaux, J. A simultaneous determination of
2585 nucleation and growth rates from batch spontaneous precipitation. *Chemical Engineering*
2586 *Science* **1999**, *54* (9), 1161-1180.
2587 (343) Nagy, Z. K.; Fujiwara, M.; Woo, X. Y.; Braatz, R. D. Determination of the Kinetic
2588 Parameters for the Crystallization of Paracetamol from Water Using Metastable Zone Width
2589 Experiments. *Industrial & Engineering Chemistry Research* **2008**, *47* (4), 1245-1252.
2590 (344) Nancollas, G. H. Kinetics of crystal growth from solution. *Journal of Crystal Growth*
2591 **1968**, *3-4*, 335-339.
2592 (345) Zipp, G. L.; Rodriguez-Hornedo, N. Determination of crystal growth kinetics from
2593 desupersaturation measurements. *International Journal of Pharmaceutics* **1989**, *51* (2), 147-156.
2594 (346) Mangood, A.; Malkaj, P.; Dalas, E. Hydroxyapatite crystallization in the presence of
2595 acetaminophen. *Journal of Crystal Growth* **2006**, *290* (2), 565-570.

(347) Kougoulos, E.; Jones, A. G.; Wood-Kaczmar, M. W. Estimation of crystallization kinetics for an organic fine chemical using a modified continuous cooling mixed suspension mixed product removal (MSMPR) crystallizer. *Journal of Crystal Growth* **2005**, *273* (3), 520-528.

(348) Khanna, V. Digital Design of Crystals: Predicting Driving Forces For Crystallization Using Atomistic Simulations. PhD, University of California Santa Barbara, 2022.

(349) Cole, K. P.; Groh, J. M.; Johnson, M. D.; Burcham, C. L.; Campbell, B. M.; Diseroad, W. D.; Heller, M. R.; Howell, J. R.; Kallman, N. J.; Koenig, T. M.; et al. Kilogram-scale prexasertib monolactate monohydrate synthesis under continuous-flow CGMP conditions. *Science* **2017**, *356* (6343), 1144-1150.

(350) Johnson, M. D.; Burcham, C. L.; May, S. A.; Calvin, J. R.; McClary Groh, J.; Myers, S. S.; Webster, L. P.; Roberts, J. C.; Reddy, V. R.; Luciani, C. V.; et al. API Continuous Cooling and Antisolvent Crystallization for Kinetic Impurity Rejection in cGMP Manufacturing. *Organic Process Research & Development* **2021**, *25* (6), 1284-1351.

(351) Polster, C. S.; Cole, K. P.; Burcham, C. L.; Campbell, B. M.; Frederick, A. L.; Hansen, M. M.; Harding, M.; Heller, M. R.; Miller, M. T.; Phillips, J. L.; et al. Pilot-Scale Continuous Production of LY2886721: Amide Formation and Reactive Crystallization. *Organic Process Research & Development* **2014**, *18* (11), 1295-1309.

(352) Szilagyi, B.; Pal, K.; Beheshti Tabar, I.; Nagy, Z. K. A Novel Robust Digital Design of a Network of Industrial Continuous Cooling Crystallizers of Dextrose Monohydrate: From Laboratory Experiments to Industrial Application. *Industrial & Engineering Chemistry Research* **2020**, *59* (51), 22231-22246.

(353) Evans, T. W.; Margolis, G.; Sarofim, A. F. Mechanisms of secondary nucleation in agitated crystallizers. *AIChE Journal* **1974**, *20* (5), 950-958.

(354) Wu, W.-L.; Oliva, J. A.; Kshirsagar, S.; Burcham, C. L.; Nagy, Z. K. Continuous In Situ Seed Generation through the Integration of a Mixed Suspension Mixed Product Removal and an Oscillatory Baffled Crystallizer for the Control of Crystal Size Distribution and Polymorphic Form. *Crystal Growth & Design* **2021**, *21* (12), 6684-6696.

(355) Mullin, J. W.; Nývlt, J. Programmed cooling of batch crystallizers. *Chemical Engineering Science* **1971**, *26* (3), 369-377.

(356) Ward, J. D.; Yu, C.-C.; Doherty, M. F. A new framework and a simpler method for the development of batch crystallization recipes. *AIChE Journal* **2011**, *57* (3), 606-617.

(357) Erratum. *AIChE Journal* **2012**, *58* (4), 1311-1311.

(358) Botsaris, G. D.; Mason, E. A.; Reid, R. C. Growth of Potassium Chloride Crystals from Aqueous Solutions. I. The Effect of Lead Chloride. *The Journal of Chemical Physics* **2004**, *45* (6), 1893-1899.

(359) Dowling, R.; Davey, R. J.; Curtis, R. A.; Han, G. J.; Poornachary, S. K.; Chow, P. S.; Tan, R. B. H. Acceleration of crystal growth rates: an unexpected effect of tailor-made additives. *Chemical Communications* **2010**, *46* (32), 5924-5926.

(360) Li, L.; Lechuga-Ballesteros, D.; Szkudlarek, B. A.; Rodríguez-Hornedo, N. r. The Effect of Additives on Glycine Crystal Growth Kinetics. *Journal of Colloid and Interface Science* **1994**, *168* (1), 8-14.

(361) Piana, S.; Jones, F.; Gale, J. D. Aspartic acid as a crystal growth catalyst. *CrystEngComm* **2007**, *9* (12), 1187-1191.

(362) Michaels, A. S.; Colville, A. R., Jr. THE EFFECT OF SURFACE ACTIVE AGENTS ON CRYSTAL GROWTH RATE AND CRYSTAL HABIT. *The Journal of Physical Chemistry* **1960**, *64* (1), 13-19.

(363) Shekunov, B. Y.; Grant, D. J. W.; Latham, R. J.; Sherwood, J. N. In Situ Optical Interferometric Studies of the Growth and Dissolution Behavior of Paracetamol (Acetaminophen) Crystals. 3. Influence of Growth in the Presence of p -Acetoxyacetanilide. *Journal of Physical Chemistry B* **1997**, *101* (44), 9107-9112.

(364) Ma, W.; Lutsko, J. F.; Rimer, J. D.; Vekilov, P. G. Antagonistic cooperativity between crystal growth modifiers. *Nature* **2020**, *577* (7791), 497-501.

(365) Kuvadia, Z. B.; Doherty, M. F. Effect of Structurally Similar Additives on Crystal Habit of Organic Molecular Crystals at Low Supersaturation. *Crystal Growth & Design* **2013**, *13* (4), 1412-1428.

(366) Lutsko, J. F.; González-Segredo, N.; Durán-Olivencia, M. A.; Maes, D.; Van Driessche, A. E. S.; Sleutel, M. Crystal Growth Cessation Revisited: The Physical Basis of Step Pinning. *Crystal Growth & Design* **2014**, *14* (11), 6129-6134.

(367) Shtukenberg, A. G.; Ward, M. D.; Kahr, B. Crystal growth inhibition by impurity stoppers, now. *Journal of Crystal Growth* **2022**, *597*, 126839.

(368) Mazal, T.; Doherty, M. F. Modeling Impurity-Mediated Crystal Growth and Morphologies of Centrosymmetric Molecules. *Crystal Growth & Design* **2023**, *23* (1), 369-379.

(369) Agrawal, P.; Rawal, S. H.; Reddy, V. R.; Viswanath, S. K.; Merritt, J. M. Case Studies in the Application of a Workflow-Based Crystallization Design for Optimized Impurity Rejection in Pharmaceutical Development. *Organic Process Research & Development* **2023**, *27* (4), 610-626.

(370) Nordstrom, F. L.; Sirota, E.; Hartmanshenn, C.; Kwok, T. T.; Paolello, M.; Li, H.; Abeyta, V.; Bramante, T.; Madrigal, E.; Behre, T.; et al. Prevalence of Impurity Retention Mechanisms in Pharmaceutical Crystallizations. *Organic Process Research & Development* **2023**, *27* (4), 723-741.

(371) Urwin, S. J.; Levilain, G.; Marziano, I.; Merritt, J. M.; Houson, I.; Ter Horst, J. H. A Structured Approach To Cope with Impurities during Industrial Crystallization Development. *Organic Process Research & Development* **2020**, *24* (8), 1443-1456.

(372) Shiea, M.; Buffo, A.; Vanni, M.; Marchisio, D. Numerical Methods for the Solution of Population Balance Equations Coupled with Computational Fluid Dynamics. *Annual Review of Chemical and Biomolecular Engineering* **2020**, *11* (1), 339-366.

(373) Fox, R. O. *Computational Models for Turbulent Reacting Flows*; Cambridge University Press, 2003. DOI: DOI: 10.1017/CBO9780511610103.

(374) Fox, R. O. Large-Eddy-Simulation Tools for Multiphase Flows. *Annual Review of Fluid Mechanics* **2012**, *44* (1), 47-76.

(375) Achermann, R.; Adams, R.; Prasser, H.-M.; Mazzotti, M. Characterization of a small-scale crystallizer using CFD simulations and X-ray CT measurements. *Chemical Engineering Science* **2022**, *256*, 117697.

(376) Zhang, M.; Liang, Z. Z.; Wu, F.; Chen, J. F.; Xue, C. Y.; Zhao, H. Crystal engineering of ibuprofen compounds: From molecule to crystal structure to morphology prediction by computational simulation and experimental study. *Journal of Crystal Growth* **2017**, *467*, 47-53.

(377) van der Lee, A.; Dumitrescu, D. G. Thermal expansion properties of organic crystals: a CSD study. *Chemical Science* **2021**, *12* (24), 8537-8547.

(378) Kriz, K.; Schmidt, L.; Andersson, A. T.; Walz, M. M.; van der Spoel, D. An Imbalance in the Force: The Need for Standardized Benchmarks for Molecular Simulation. *Journal of Chemical Information and Modeling* **2023**, *63* (2), 412-431.

(379) James, A.; John, C.; Melekamburath, A.; Rajeevan, M.; Swathi, R. S. A journey toward the heaven of chemical fidelity of intermolecular force fields. *WIREs Computational Molecular Science* **2022**, *12* (4), e1599.

(380) Uzoh, O. G.; Galek, P. T. A.; Price, S. L. Analysis of the conformational profiles of fenamates shows route towards novel, higher accuracy, force-fields for pharmaceuticals. *Physical Chemistry and Chemical Physics* **2015**, *17*, 7936-7948.

(381) Maurer, R. J.; Freysoldt, C.; Reilly, A. M.; Brandenburg, J. G.; Hofmann, O. T.; Bjorkman, T.; Lebegue, S.; Tkatchenko, A.; Clarke, D. R. Advances in Density-Functional Calculations for Materials Modeling. *Annual Review of Materials Research, Vol 49* **2019**, *49*, 1-30.

(382) Hunnisett, L. M.; al., e. The Seventh Blind Test of Crystal Structure Prediction: Structure Generation Methods. *Acta Crystallographica Section B, Structural Science, Crystal Engineering and Materials* **2024**, *in preparation*.

(383) Hunnisett, L. M.; al, e. The Seventh Blind Test of Crystal Structure Prediction: Structure Ranking Methods. *Acta Crystallographica Section B, Structural Science, Crystal Engineering and Materials* **2024**, *in preparation*.

(384) Clements, R. J.; Dickman, J.; Johal, J.; Martin, J.; Glover, J.; Day, G. M. Roles and opportunities for machine learning in organic molecular crystal structure prediction and its applications. *MRS Bulletin* **2022**, *47* (10), 1054-1062.

(385) Price, L. S.; Price, S. L. Packing Preferences of Chalcones: A Model Conjugated Pharmaceutical Scaffold. *Crystal Growth & Design* **2022**, *22* (3), 1801-1816.

(386) Bhardwaj, R. M.; McMahon, J. A.; Nyman, J.; Price, L. S.; Konar, S.; Oswald, I. D. H.; Pulham, C. R.; Price, S. L.; Reutzel-Edens, S. M. A Prolific Solvate Former, Galunisertib, under the Pressure of Crystal Structure Prediction, Produces Ten Diverse Polymorphs. *Journal of the American Chemical Society* **2019**, *141* (35), 13887-13897.

(387) Braun, D. E.; Oberarcher, H.; Arnhard, K.; Orlova, M.; Griesser, U. J. 4-Aminoquinaldine monohydrate polymorphism: Prediction and impurity aided discovery of a difficult to access stable form. *CrystEngComm* **2016**, *18*, 4053-4067.

(388) Braun, D. E.; Vickers, M.; Griesser, U. J. Dapsone Form V: A Late Appearing Thermodynamic Polymorph of a Pharmaceutical. *Molecular Pharmaceutics* **2019**, *16* (7), 3221-3236.

(389) Burcham, C. L. Crystalline compound and a process for its preparation. US 8,299,059, 2012.

2721   For Table of Contents Use Only

2722