Acc. Chem. Res. **1994,** *27,* 309−314

Are Crystal Structures Predictable?

Angelo Gavezzotti*

Dipartimento di Chimica Fisica ed Elettrochimica, Università di Milano, Milano, Italy

Received May 16, 1994

"No": by just writing down this concise statement, in what would be the first one-word paper in the chemical literature, one could safely summarize the present state of affairs, earn an honorarium from the American Chemical Society, and do a reasonably good service to his or her own reputation. In the mainstream of academic tradition, one could then concede a "maybe", or even a conditional "yes", thus making a good point for discussion; and then, in the mainstream of publication policy tradition, proceed eventually to have his or her papers rejected by referees taking the opposite stand.

Fortunately, there is a rhetorical way out of this predicament, known to medieval philosophers as *amplificatio*: in plain words it means, when you cannot provide an answer, just rephrase and expand the statement of the question. To this very old trick we will resort in this paper. In fact, the title question is a bit too straightforward and simple-minded; such broad terms as "crystal structure" and "prediction" need be defined in more detail. There are several levels of desirable *a priori* information on a solid; they will be described by posing a number of typical, more restricted questions, in order of increasing complexity. Organic substances only will be considered.

It is assumed that it need not be explained to the reader why control or prediction of the structure of a solid, at a molecular level, is desirable; there are several self-evident justifications, on both theoretical and practical grounds, for striving to understand the basic factors that dictate the arrangement of molecules in space when they recognize each other at a short distance and eventually coagulate in a rigid configuration. While the present knowledge of intramolecular valence can be considered satisfactory, that of intermolecular "valence" is rudimentary; and the perspective of being able to design molecular solids with predetermined physical properties, which depend on structure, is appealing (an understatement) to applied chemists in the fields of pigments,[1] pharmaceuticals,[2] magnets,[3] conductors,[4] and photosensitive[5] or optoelectronic[6] materials. So one has here a big theoretical challenge going hand in hand with big business.

In the early days of X-ray crystallography, guessing at the crystal structure by minimizing intermolecular repulsions was considered a viable method of solving the phase problem, when cell dimensions and diffraction intensities were available. From such a perspective, knowledge of the cell volume implied that intermolecular attractions had been satisfied, and that only mutual avoidance between rigid objects had to be accomplished, either by rough (but surprisingly efficient) mechanical devices[7a] or by computer sieving.[7b] These procedures were suddenly made obsolete, and dismissed, by the advent of direct methods. Crystal structure prediction resurfaced only in very recent times, and with a much more ambitious connotation; the new problem is to consider an organic compound for which a structural formula has been written on paper, but whose synthesis (presumably expensive in terms of materials or human resources) has not yet been accomplished. In keeping with the rhetorical profile of this paper, typical questions on its future as a solid will now be posed.

1. *Will this compound crystallize at all*? Thermodynamics holds that any substance must crystallize, provided it is pure and the temperature is low (or pressure is high) enough. But organic chemistry thrives in mild temperature−pressure regimes, prone to the much more elusive dictates of kinetics. Dissolution always works in the proper solvent while crystal growth from solution is problematic; melting nearly always occurs at higher temperatures than freezing; a crystal is more readily destroyed than built. The organic solid state ranges from waxes or glasses to disordered, strained, or twinned crystals, to powders, and eventually, to well-shaped single crystals. Chemists often come to grips with tough problems in the control of solidification, crystal growth, and crystal morphology, mainly due to the perverse kinetic control of nucleation; and this is a well-developed research field of its own.[8]

For example, sexithienyl, a compound of great importance in nonlinear optics, has a high melting point, yet no single crystals of this substance could be grown, in spite of considerable effort. A reasonable and stable crystal structure has been predicted[9] by calculations based on empirical potentials. Recently, a Rietveld analysis of powder specimens (the best that

Angelo Gavezzotti (born in 1944) studied chemistry at the University of Milano and graduated in 1968 with a thesis in X-ray crystallography. He has worked in the fields of theoretical and structural chemistry with supervision and friendly advice from M. Simonetta. He spent research terms in Orsay, France (1973), and in Ann Arbor, MI (1977–1978), with L. S. Bartell. He is presently professor of physical chemistry at the University of Milano, and his research interests focus on the structural chemistry of organic crystals. He served a three-year term as Coeditor of *Acta Crystallographica*.

---

* New address: Dipartimento di Chimica Strutturale e Stereochimica Inorganica, Università di Milano, Milano, Italy.

(1) Klebe, G.; Graser, F.; Hädicke, E.; Berndt, J. *Acta Crystallogr.* **1989**, *B45*, 69−77.

(2) Haleblian, J.; McCrone, W. *J. Pharm. Sci.* **1969**, *58*, 911−929.

(3) Miller, J. S.; Epstein, A. J.; Reiff, W. M. *Acc. Chem. Res.* **1988**, *21*, 114−120.

(4) Hunig, S.; Erk, P. *Adv. Mater.* **1991**, *3*, 225−236.

(5) For the structural problems connected with epitaxy and photoconductivity of small molecules on polymers, see: Scaringe, R. P.; Perez, S. *J. Phys. Chem.* **1987**, *91*, 2394−2403 and references therein.

(6) Chemla, D. S.; Zyss, J. *Nonlinear Optical Properties of Organic Molecules and Crystals*; Academic Press: Orlando, 1987.

(7) (a) Kitaigorodski, A. I. *Organic Chemical Crystallography*; Consultants Bureau: New York, 1961 (the structure-seeking apparatus). For a review of Kitaigorodski's work, see also: Struchkov, Y. T.; Fedin, E. I. *Acta Chem. Hung.* **1993**, *130*, 159−172. (b) Rabinovich, D.; Schmidt, G. M. J. *Nature (London)* **1966**, *211*, 1391−1393.

(8) Hulliger, J. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 143−162.

(9) Gavezzotti, A.; Filippini, G. *Synth. Met.* **1991**, *40*, 257−266.

0001-4842/94/0127-0309$04.50/0

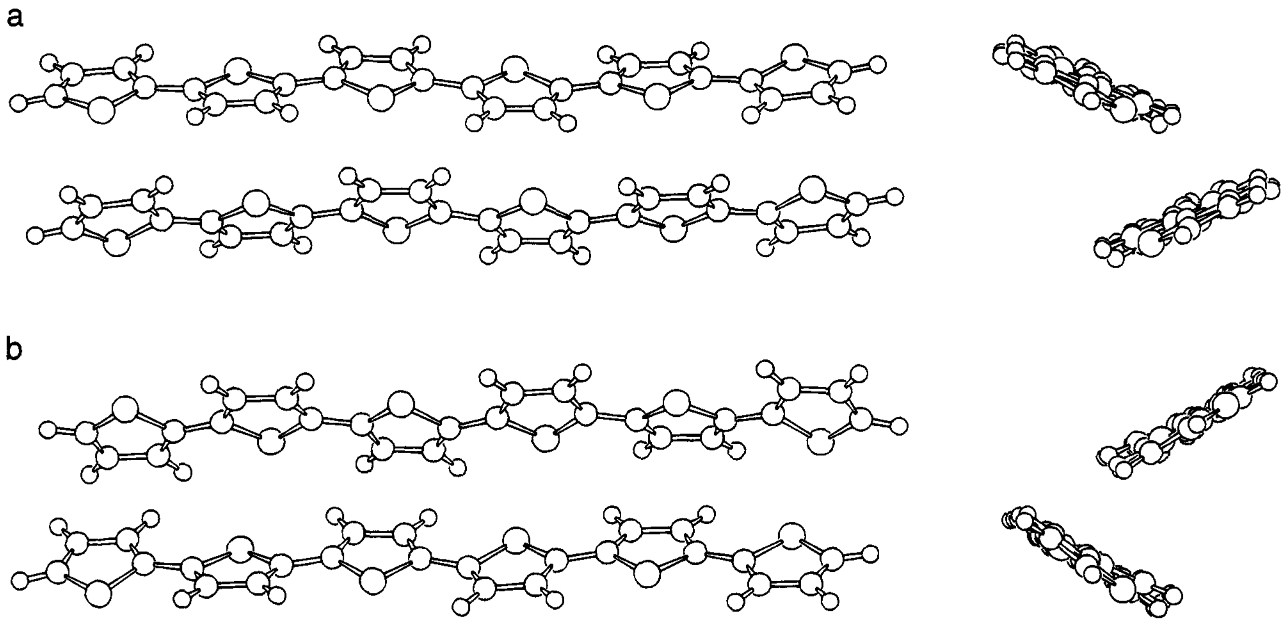

Figure 1. (a) The main motif in the predicted crystal structure of sexithienyl (ref 9; $P2_1/a$, $Z = 2$). (b) The same for the structure from a Rietveld refinement of powder data (ref 10; $P2_1/c$, $Z = 4$). The two structures differ mainly in the interplanar angle between neighbor molecules (49° vs 67°), better shown in the side views on the right.

could be obtained) has been published.[10] While the agreement between the main features of the predicted and experimental crystal structures is pleasing (Figure 1), the riddle of the lack of sexithienyl single crystals is still unanswered.

2. *Is this crystal high-melting*? The melting temperature (T_m) is high for high melting enthalpy or for low melting entropy. The entropic factor implies that disordered crystals, or crystals whose liquids are heavily associated (*e.g.*, by hydrogen bonding), have higher T_m's. Correlations between T_m and crystal cohesion should therefore be taken with caution.

A very old rule of thumb states that more symmetric molecules form higher-melting crystals;[11] this idea has been analyzed[12] using ortho-, meta-, and para-disubstituted benzenes (XC_6H_4Y, X and Y being any substituents). A survey of their T_m's shows that para isomers are the highest-melting ones, with very few exceptions; for only 18 out of 238 para–meta and para–ortho couples, the para isomer melts at a lower temperature. However, the definition of molecular symmetry in this context is really elusive and merges uncomfortably with that of molecular shape. The rule of thumb stays such, and cannot be given a sound theoretical or structural foundation. T_m is still one of the most difficult crystal properties to predict.

3. *What is the lattice energy* (*heat of sublimation*)? Extensive statistical studies have been conducted on relationships between molecular and crystal properties for non-hydrogen-bonding compounds containing C, H, N, O, S, and Cl atoms,[13-16] as well as for the most common families of hydrogen-bonding compounds[17] (acids, alcohols, and amides). Correlations were found which allow an estimate of sublimation enthalpies from molecular parameters like the number of valence electrons (Z) or the van der Waals surface (S). For example, in non-hydrogen-bonded oxohydrocarbons,

$$\Delta H_s = 0.201Z + 9.4 \text{ kcal/mol}$$

$$\Delta H_s = 0.077S(\text{Å}^2) + 8.9 \text{ kcal/mol}$$

Standard deviations of these linear regressions are comparable to experimental uncertainties of measurements;[18] at least in this respect, truly predictive correlations between molecular and crystal properties can be established. In some cases, errors in experimental ΔH_s's have been detected by redeterminations prompted by large deviations from the correlation.[14] Needless to say, the total lattice energy as such carries no information on the geometrical structure of the crystal.

4. *Will the crystal structure be non-centrosymmetric*? This is a simple but vital requirement for some practical applications of crystal chemistry.[19] Crystal centrosymmetry is often a matter of debate, and it is sometimes one of the refinable parameters in X-ray crystal structure analysis, rather than a stringent *a priori* condition.[20] One sees here a wide gap between the high (sometimes too high) resolution of diffraction experiments, where a single non-centrosymmetrically arranged atom in a large molecule would make a total difference, and the coarse view of the applied chemist. No one, except a neutron diffractionist, would consider non-centrosymmetric a hypothetical $P2_1$ crystal structure of monodeuteriobenzene.

---

[10] Porzio, W.; Destri, S.; Mascherpa, M.; Bruckner, S. *Acta Polym.* **1993**, *44*, 266–272.

[11] An early statement is by Hückel: Hückel, W. *Theoretische Grundlage der Organischen Chemie*; Akademische Verlagsgesellschaft: Leipzig, 1931; Vol. II, pp 185–186.

[12] Gavezzotti, A. To be published.

[13] Gavezzotti, A. *J. Am. Chem. Soc.* **1989**, *111*, 1835–1843.

[14] Gavezzotti, A. *J. Phys. Chem.* **1991**, *95*, 8948–8955.

[15] Gavezzotti, A.; Filippini, G. *Acta Crystallogr.* **1992**, *B48*, 537–545.

[16] Gavezzotti, A.; Filippini, G. *Acta Chim. Hung.* **1993**, *130*, 205–220.

[17] Gavezzotti, A.; Filippini, G. *J. Phys. Chem.* **1994**, *98*, 4831–4837.

[18] For a review of available sublimation enthalpies of organic compounds, see: Chickos, J. S. In *Molecular Structure and Energetics*; Liebman, J. F., Greenberg, A., Eds.; VCH: New York, 1987; Vol. 2.

[19] Paul, I. C.; Curtin, D. Y. *Chem. Rev.* **1981**, *81*, 525–541.

[20] For a $P1$ reassigned as $P\bar{1}$, see: Marsh, R. E. *Acta Crystallogr.* **1990**, *C46*, 1356–1357. See also: Marsh, R. E. *Acta Crystallogr.* **1994**, *A50*, 450, 455. The author is so assiduous in this kind of exercise that papers so reconsidered are commonly said to have been "marshed".

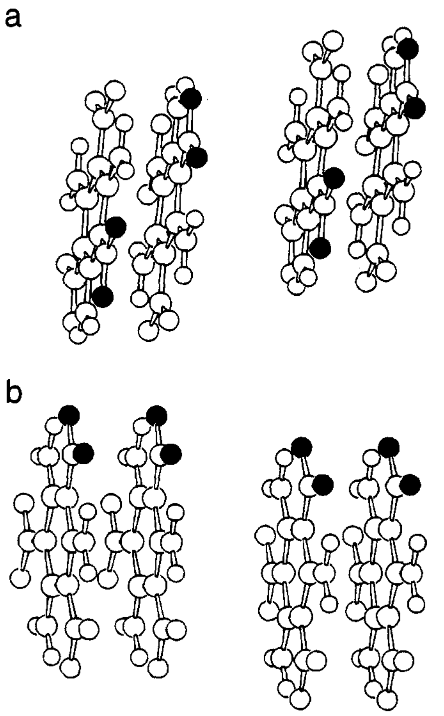

Figure 2. Arrangement of molecules in (a) the X-ray crystal structure of 1,3,5-triamino-2,4,6-trinitrobenzene (ref 24; $P\bar{1}$, $Z = 2$) and (b) the simulated crystal structure (ref 26; $P1$, $Z = 2$). Oxygen atoms in one nitro group are filled in.

The opinion that molecules with a high dipole moment tend to crystallize in a head-to-tail centrosymmetric fashion is untenable, as has been demonstrated by a detailed analysis:[21] the dipole representation of a charge distribution applies at large distances from it, while neighbor molecules in crystals see each other at distances comparable to molecular dimensions. On the other hand, the carboxylic acid group nearly always forces crystal centrosymmetry by forming cyclic dimers.[17,22] As is often the case, we only know how to produce the effect we do not want.

A crystal grown out of a solution containing only one enantiomer will perforce be non-centrosymmetric, but nothing can be said *a priori* on the spontaneous resolution of racemic solutions by crystallization. The relative stability of resolved and racemic crystals has been analyzed,[23] but there are at present no really predictive concepts on this fascinating subject, which may be related to the chirality of the chemistry of life.

Quite often, non-centrosymmetric molecular layers are readily formed, but they cannot be prevented from assuming an apparently very favorable centrosymmetric arrangement in the crystal. For example, the crystal structure of 1,3,5-triamino-2,4,6-trinitrobenzene has been assigned to a centrosymmetric space group ($P\bar{1}$) by X-ray analysis,[24] while the material displays a second harmonic generation propensity,[25] a property of non-centrosymmetric structures. Plausible non-centrosymmetric structures, with lattice energies quite comparable to that of the X-ray one, have been generated (Figure 2); the discussion of the results[26] has a lot of academic ifs and buts, perhaps contributing to confusion more than to the advancement of knowledge. The formation of non-centrosymmetric domains seems, however, the most likely explanation of the unusual properties of this crystal.

5. *Will some parts of the molecule take up a predictable orientation in the crystal?* Use of the information contained in the Cambridge Structural Database[27] has led to a number of statistical studies on the geometry of hydrogen bonding, of halogen–halogen interactions, and of other preferred approach paths between chemically recognizable molecular moieties. The reader is referred to an excellent review[28] on the subject.

Much work (and speculation) has been devoted[29] to the so-called $\pi-\pi$ interactions between aromatic rings, driving to stacking, against the "electrostatic" attractions between rim hydrogen atoms and core carbon atoms, driving to T-shaped arrangements; preference for the latter is often assumed, quoting as a key example the benzene crystal, which in fact does contain also almost stacked neighbor molecules. A paper[30] in which the distribution of phenyl group orientations in hydrocarbon crystals has been examined, with peaks for both parallel and T-shaped arrangements, and a non-negligible population in between, has not been considered too seriously. Rules for the prediction of the appearance of herringbone *versus* stacked motifs in condensed aromatics have, apparently, been derived.[31]

In crystals of *monofunctional* carboxylic acids and amides, virtually no exceptions to the formation of cyclic dimers for the former and of single N–H···O=C hydrogen bonds in the latter were found.[17] Hydrogen bond formation has undoubtedly a very high priority in the construction of a crystal structure, but molecules with several acceptor and/or donor groups quite often crystallize in different polymorphic forms with different hydrogen-bonding networks.[32]

To conclude this section, one could say that some broad trends in the dependence of crystal packing from the presence of certain substituents or fragments have been identified; but this "substituent effect" in crystal chemistry stands on a shaky pedestal, since interactions in crystals of complex molecules are diverse and diffuse, and relying on local effects is always dangerous.

6. *What can be the space group and the number of molecules in the asymmetric unit?* The very concept of "space group" needs a little revision for crystal chemistry purposes. The presence or absence of a center of symmetry may be questionable;[20] the same applies to every symmetry element. To the eyes of an X-ray crystallographer, a glide plane is or is not present according to an extinction pattern, but the borderline between extinct and very weak reflections can sometimes be a matter of subjective judgement (parasitic diffraction phenomena also contribute). Minor molecular displacements may destroy some

---

(21) Whitesell, J. K.; Davis, R. E.; Saunders, L. L.; Wilson, R. J.; Feagins, J. P. *J. Am. Chem. Soc.* **1991**, *113*, 3267–3270.

(22) Leiserowitz, L. *Acta Crystallogr.* **1976**, *B32*, 775–802.

(23) Brock, C. P.; Schweizer, W. B.; Dunitz, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 9811–9820.

(24) Cady, H. H.; Larson, A. C. *Acta Crystallogr.* **1965**, *18*, 485–496.

(25) Ledoux, I.; Zyss, J.; Siegel, J. S.; Brienne, J.; Lehn, J. M. *Chem. Phys. Lett.* **1990**, *172*, 440–444.

(26) Filippini, G.; Gavezzotti, A. Submitted.

(27) Allen, F. H.; Kennard, O.; Taylor, R. *Acc. Chem. Res.* **1983**, *16*, 146–153.

(28) Desiraju, G. R. *Crystal Engineering*; Elsevier: Amsterdam, 1989.

(29) See, *e.g.*: Dahl, T. *Acta Chem. Scand.* **1994**, *48*, 95–106 and references therein.

(30) Gavezzotti, A. *Chem. Phys. Lett.* **1989**, *161*, 67–72.

(31) Desiraju, G. R.; Gavezzotti, A. *Acta Crystallogr.* **1989**, *B45*, 473–482.

(32) Sulfa drugs provide striking examples: see, *e.g.*: Bar, I.; Bernstein, J. *J. Pharm. Sci.* **1985**, *74*, 255–263.

symmetry element and bring about a change in space group (to the overdetailed eyes of the X-ray analyst), without really affecting the properties of the solid. For the crystal chemist, the prediction of the space group may be a whimsical exercise, if what counts is just a broad understanding of how molecules arrange themselves in space. Besides, in a molecular crystal (here meaning one in which distinguishable chemical entities appear, for which forces within the entity are considerably stronger than forces between entities) a distinction must be made between intramolecular, or point-group, symmetry and the true "intermolecular" symmetry, when the asymmetric unit is less than one molecule.[33]

Overall, crystal symmetry has two facets. On one side, in a milestone *mathematical* development, it was demonstrated that the combinations of symmetry elements give rise to no fewer and no more than 230 independent three-dimensional space groups. On the other side, crystal symmetry has to do with the mutual recognition of molecules to form a stable solid, a fascinating and essentially *chemical* problem that requires an evaluation of intermolecular forces. It should be clear that no necessary relationship holds between these two views; 230 space groups exist, but molecules cannot freely choose among them. Far from it, there are rather strict conditions that can be met only by a limited number of combinations of very few symmetry elements; for organic compounds, these are the inversion center, the 2-fold screw axis, and the glide plane, plus the ubiquitous translation (sometimes disguised as centering), itself a respectable, if often forgotten, symmetry operator. Thus, the choice of the space group for organic crystals is usually restricted to those including the above combinations: $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$, $C2/c$, $P2_12_12_1$, $Pbca$. The well-known statistics on space group populations[34] for organic compounds confirms this, as Kitaigorodski pointed out decades ago.[35]

Some crystals reach a stable (or at least a lasting metastable) state with more than one molecule in the asymmetric unit. Statistics on the Cambridge Database have these occurrences at 8.3%,[36] but this is presumably an underestimation, since the Database is socially biased: structures with several molecules in the asymmetric unit pose a small supplementary technical problem in final space group assignment and structure refinement and were often in the past (and probably still are) put aside by busy crystallographers as unsavory members of their waiting lists. Once again, the reader is reminded of the discussion on the presence or absence of a symmetry operator, in this case the one that could provide a relationship between the partners of the plurimolecular asymmetric unit.

Some basic rules that preside over the formation of intra- and intermolecular hydrogen bonding have been identified.[37] In addition, it turns out that molecules which form very stable clusters in the liquid by hydrogen bonding are more likely to form plurimolecular asymmetric units, since these clusters are carried over intact into the crystal, and perfect symmetry within them is energetically irrelevant, or even slightly unfavorable: 40% of the alcohol crystals in the Cambridge Database have more than one molecule in the asymmetric unit.[17] For non-hydrogen-bonded crystals a similar explanation may be proposed, although no simple rules based on chemical reasoning can be put forward for preaggregation in the liquid state.

7. *What are the cell parameters?* The cell volume per molecule is rather easily estimated from molecular volume, after the Kitaigorodski idea of a constant packing coefficient;[35] hence, the crystal density too can be roughly estimated (see refs 15 and 17 for average packing coefficients of different chemical classes). If space is to be efficiently used in a condensed phase, there must be broad correlations between molecular dimensions and cell edges: for example, if D_s is the shortest molecular dimension, C_s the shortest cell edge, D_h the longest molecular dimension, and C_h the longest cell edge, the following restrictions apply[38] (Å):

$$D_s - 2 < C_s < D_s + 5$$

$$C_h > D_h - 3$$

Cell dimensions are indeed a bad identifier of a crystal structure, since their choice is not always unique. Distances between molecular centers of mass may be more useful; of course, some of these coincide with the length of screw or glide translations and, hence, are equal to one-half the cell parameters along unique crystallographic axes. These distances are the main quantitative descriptors of crystal geometry and are dictated solely by the strength and directionality of intermolecular forces. At this level, therefore, not much can be predicted with decent accuracy unless quantitative intermolecular potentials are available.

The systematic calibration of a set of potential energy parameters for organic crystals containing H, C, N, O, S, and Cl atoms, with[17] or without[39] hydrogen bonds, has been (painstakingly) accomplished. The reader will be spared the details of, and the endless disputes on, the methods employed in such work; space forbids also a quotation of the many alternative force fields available in the literature.[40] Suffice it to say that these parameters are as few as possible, and that with them one can safely calculate lattice energies (since experimental heats of sublimation[18] are reproduced), trusting that lattice dynamics is not grossly misrepresented (since reasonable lattice vibration frequencies are calculated[41] for observed crystal structures). The functional form includes one exponential and one inverse sixth power term in interatomic distances, so that computing times are not inflated by slowly-converging summations. These potentials (Table 1) have been tailored for the explicit task of performing large scale searches of crystal potential surfaces, or, in fewer words, for crystal structure prediction.

8. *Are crystal structures predictable?* Of course, the final question is whether it is possible to predict *ab initio* the complete structure of any organic crystal,

---

[33] See the discussion in the following: Scaringe, R. P. In *Electron Crystallography of Organic Molecules*; Fryer, J. R., Dorset, D. L., Eds.; Kluwer: Dordrecht, 1991, especially pp 92–94.

[34] Baur, W. H.; Kassner, D. *Acta Crystallogr.* **1992**, *B48*, 356–369.

[35] See ref 7a, introductory chapters.

[36] Padmaja, N.; Ramakumar, S.; Viswamitra, M. A. *Acta Crystallogr.* **1990**, *A46*, 725–730.

[37] Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120–126.

[38] Gavezzotti, A. *J. Am. Chem. Soc.* **1991**, *113*, 4622–4629.

[39] Filippini, G.; Gavezzotti, A. *Acta Crystallogr.* **1993**, *B49*, 868–880.

[40] See refs 17 and 39 for perspective and discussion; see also: Pertsin, A. J.; Kitaigorodski, A. I. *The Atom–Atom Potential Method*; Springer-Verlag: Berlin, 1987.

[41] The lattice-dynamical procedure is described in the following: Filippini, G.; Gramaccioli, C. M. *Acta Crystallogr.* **1986**, *B42*, 605–609.

Table 1. Atom–Atom Potential Parameters: $E = A\ \exp(-BR_{ij}) - CR_{ij}^{-6}$

interaction	A^a	B^b	C^c	ϵ^d	R^e
H···H	5774	4.01	26.1	0.010	3.36
H···C	28 870	4.10	113	0.049	3.29
H···N	54 560	4.52	120	0.094	2.99
H···O	70 610	4.82	105	0.121	2.80
H···S	64 190	4.03	279	0.110	3.35
H···Cl	70 020	4.09	279	0.120	3.30
C···C	54 050	3.47	578	0.093	3.89
C···N	117 470	3.86	667	0.201	3.50
C···O	93 950	3.74	641	0.161	3.61
C···S	126 460	3.41	1504	0.217	3.96
C···Cl	93 370	3.52	923	0.160	3.83
N···N	87 300	3.65	691	0.150	3.70
N···O	64 190	3.86	364	0.110	3.50
O···O	46 680	3.74	319	0.080	3.61
O···S	110 160	3.63	906	0.189	3.72
O···Cl	80 855	3.63	665	0.139	3.72
S···S	259 960	3.52	2571	0.445	3.83
Cl···Cl	140 050	3.52	1385	0.240	3.83
HB···O (amides)	3 607 810	7.78	238	4.0	1.80
HB···O (acids)	6 313 670	8.75	205	7.0	1.60
HB···O (alcohols)	4 509 750	7.78	298	5.0	1.80
HB···N (=N–H···N)	7 215 600	7.78	476	8.0	1.80
HB···N (–NH_2···N)	1 803 920	7.37	165	2.0	1.90

[a] Kcal/mol. [b] $Å^{-1}$. [c] Kcal/(mol·$Å^{-6}$). [d] Potential well depth (kcal/mol). [e] Distance at the minimum (Å). From refs 17 and 39.

space group, cell parameters, and atomic positions, much in the same style as in X-ray single-crystal structure analysis. The answer here is definitely "no".

Undoubtedly, no true prediction in the above sense can be accomplished without calculating the crystal potential energy, but one fundamental point is the choice of the best coordinates for the energy space. Intramolecular structure can be described by just a few (mostly torsional) conformational parameters, full relaxation of intramolecular vibrational degrees of freedom being pointless, since coupling with intermolecular vibrations is negligible. The location of molecules in the cell is described by three translational and three rotational rigid-body coordinates (restrictions apply for some point-group symmetries). Rather than using space groups and cell parameters, it is more convenient[38] to start from the constituents of spacial symmetry, that is, the four basic symmetry operators (inversion center, screw, glide, and translation); molecular clusters are built under their action, and their energies are calculated by empirical potentials. In this procedure, molecular conformation must be assumed as fixed, and the fact that polymorphs may exist with different molecular conformations is one addition to an already uncomfortably long list of difficulties. Anyway, a number of promising clusters are selected and are translated in space or coupled with other operators until a full three-dimensionally periodic crystal structure is reached.[42] One advantage of this procedure is that, say, a two-molecular cluster over a center of symmetry can be used to try both $P\bar{1}$ and $P2_1/c$. The most questionable feature is that there is no guarantee that a stable cluster will actually appear in the crystal, whose stability is determined by the overall features of its three-dimensional structure.

Methods[38,43] which involve an examination of a great many possible crystal structures, using strategical shortcuts and sequential sieves, may be called "static". Their success in full prediction has been modest, but encouraging; they should be helpful when auxiliary information—from spectroscopy, powder or partial single-crystal diffraction, or structural correlation to similar compounds—is available. The construction of stable aggregates is made much easier when the consideration of predominant hydrogen-bonding schemes is possible.[44]

A "dynamic" approach uses Monte Carlo or molecular dynamics calculations.[45–47] The starting point is a collection of molecules in random orientations, and the predicted equilibrium state is the result of averaging over a large configurational space, or of evolution in time after solution of the classical equations of motion. In both cases, molecular interactions must be calculated by empirical potentials, which retain their pivotal role in the whole procedure. Computing times increase steeply with the number of molecules in the statistical sample and put a severe strain even on present-day machines. Ideally, this approach allows the simulation of the complete phase behavior of the substance, as a function of temperature and pressure. Although its scope and promise are certainly wider than those of the static approach, only the reproduction of the crystal structure of benzene[48] and a few other organic molecules[46] has been achieved so far, and a definite proof that such methods can give an unequivocal solution to the problem of crystal structure prediction has not been produced. The author of the present paper would be more than happy if this statement could be falsified in the near future. The computer software described and used in ref 46 (presumably the best available at the moment) is now being commercialized by a profit company (module "Polymorph" of the CERIUS package, by Molecular Simulations).

Polymorphism

Does the blame for the present, hardly satisfactory situation lie with technicalities? Is it just a matter of better path-finding algorithms and faster computers, or are there other basic obstacles to crystal structure prediction by calculations? There are. All computations and experiments demonstrate that many crystal structures for the same compound have quite similar lattice energies, or heats of sublimation. The ΔH_s of a medium-size organic molecule is 20–50 kcal/mol; heats of melting (ΔH_m) are typically $^1/_3$ of ΔH_s; enthalpy differences between crystalline phases[49] must be, in all evidence, just a fraction of ΔH_m, or something like 1–5 kcal/mole: just the range of experimental uncertainties of ΔH_s's on which empirical potentials are calibrated. Besides, crystalline

---

(42) Gavezzotti, A. *PROMET: A Program for the Generation of Possible Crystal Structures from the Molecular Structure of Organic Compounds*, and *Space Group Symmetry: A Primer*, University of Milano, 1993 (available from the author upon request). Using this program is an excellent way of learning the basics of space group symmetry.

(43) Holden, J. R.; Du, Z.; Ammon, H. L. *J. Comput. Chem.* **1993**, *14*, 422–437.

(44) See, for example: Zerkowski, J. A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 4298–4304.

(45) Linert, W.; Renz, F. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 776–781.

(46) Karfunkel, H. R.; Gdanitz, J. R. *J. Comput. Chem.* **1992**, *13*, 1171–1183.

(47) Perlstein, J. *J. Am. Chem. Soc.* **1994**, *116*, 455–470.

(48) Gdanitz, R. J. *Chem. Phys. Lett.* **1992**, *190*, 391–396.

(49) A vast literature deals with differences in solution enthalpies and phase transformation enthalpies of polymorphic pharmaceuticals; typical results are from a few down to fractions of a kcal/mol. For one example, see: Kojima, H.; Kiwada, H.; Kato, Y. *Chem. Pharm. Bull.* **1982**, *30*, 1824–1830. See also ref 32 and references therein.

Table 2. Organization and Tools of an Ideal Department of Organic Solid State Chemistry

experimental	theoretical
synthetic chemistry	Cambridge Structural Database
commercially available compounds	
	First-principles calculations on sample systems
(re)crystallization	
studies of nucleation	
studies of morphology	empirical methods
systematic search for polymorphs	crystal potentials
standardized studies of dissolution	
	lattice dynamics
X-ray crystallography	
single crystal	search of crystal potential hypersurfaces
powder and Rietveld	
	Monte Carlo methods
thermochemistry	molecular dynamics
differential scanning calorimetry	
thermogravimetry	
vapor pressure measurements	
solid-state NMR	
ultramicroscopy (SEM/TEM, AFM)[a]	

[a] The acronyms refer to surface visualization techniques: scanning and tunneling electron microscopy; atomic force microscopy.

phases with higher enthalpy are usually less dense and, hence, will tend to have a higher vibrational entropy; thus, free energy differences are likely to be even smaller than enthalpy differences. And, above all, thermodynamic laws are systematically violated by organic crystals; witness the frequent observation of coexisting and (even more outrageously) separately melting polymorphs. Kinetic aspects are crucial here.

Even a casual consideration of the literature on polymorphism[50] provides convincing evidence that this phenomenon is pervasive. Therefore, the problem should not be formulated as prediction of *the* crystal structure of an organic compound, but as (1) generation of a set of polymorphs, (2) unequivocal detection of the most stable one at a given temperature, and (3) modeling of the nucleation kinetics to determine which phase will actually appear under given conditions. Empirical potentials may be intrinsically too crude to handle the thermodynamic aspects of the problem, unless it is proved (as we are presently trying to do,[51] with reasonably good chances of success) that the approximations do not affect *comparisons* of lattice energies and vibrational entropies of polymorphs.[52] There is no sound and well-accepted method for treating the kinetic aspects of crystal nucleation and of phase transformations at a molecular level, although pioneering contributions, again using molecular dynamics methods, have appeared.[53]

Concluding Remarks

Organic solid state chemistry is still a sporadic discipline. Organic solid state problems are approached from different points of view, with different methods and for different purposes: vapor pressures of solids are measured for pesticides; densities for explosives; heats of fusion and solubilities for pharmaceuticals. Crystallographers just consider solids as stiff containers for the molecular systems they want to study. Prior to solution of the structure prediction problem, the principles of organic solid state chemistry should be worked out by a deliberate and coordinated effort. The ideal laboratory for this task might be organized as in Table 2: its specific purpose should be a systematic study of large classes of organic solids for the establishment of the thermodynamic, kinetic, and structural principles of the organization of matter in solids at a molecular level.[54]

In the absence of such a chimeric entity, the state of the art in organic crystal structure prediction can be summarized by saying that we are still at the stage of exploration and data collection, with a reasonable amount of data organization. The problem is, at least, being seriously considered, methods are being developed and tested, and pieces of the puzzle are slowly falling into place. A cooperation between statistical studies on available crystal structures, energetic studies based on crystal thermodynamics and kinetics, and static or dynamic computational techniques will presumably guide theoretical chemists to real predictions of crystal structure without their having to synthesize the molecule. Whether this will be for the next years or the next decades remains to be seen. The ideas and the work summarized in this Account will (hopefully) be acknowledged as preliminary but useful, and indispensable in paving the way.

Massimo Simonetta founded in the 1960s an X-ray and structural chemistry laboratory in Milano; incentives to the study of crystal packing came originally from the author's collaboration with him and other members of that group and were later corroborated by suggestions from the discussions with J. D. Dunitz. In recent times, the valuable cooperation of Giuseppe Filippini is gratefully acknowledged. This paper is dedicated to the memory of my wife, Alessandra Gara.

---

(50) See the extensive production of Kuhnert-Brandstaetter and coworkers; an entry point to this literature is the following: Kuhnert-Brandstaetter, M.; Solliger, H. W. *Mikrochim. Acta (Wien)* **1990**, *III*, 247–258.

(51) Filippini, G.; Gavezzotti, A. Work in progress.

(52) When the use of empirical methods is questionable, one should go back to first principles; but crystal structure prediction for an organic compound by quantum chemical calculations is at present, and most likely will be in the not-so-near future, impossible, since calculations without electron correlation are useless, and calculations including it are unfeasible. Carefully designed calculations on small key systems are, however, vital for a better understanding of intermolecular forces. Even a cursory mention of the vast literature in the field would be beyond the scope of this Account.

(53) (a) Bartell, L. S. *J. Phys. Chem.* **1990**, *94*, 5102–5109. (b) Chen, J.; Bartell, L. S. *J. Phys. Chem.* **1993**, *97*, 10645–10648.

(54) Such a laboratory could be set up in an academic environment with an initial investment of about 2 million dollars, and work with a yearly budget of half a million—peanuts in comparison with, say, money allocations for fancy high-energy physics machinery whose applicative spinoffs, if any, are much more problematic.