Molecular Shape and Crystal Packing: a Study of $C_{12}H_{12}$ Isomers, Real and Imaginary

by **Jack D. Dunitz***[a]), **Giuseppe Filippini**[b]) and **Angelo Gavezzotti**[c])

[a]) Organic Chemistry Laboratory, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zurich
[b]) CSRSRC CNR, c/o Dipartimento di Chimica Fisica ed Elettrochimica, Università di Milano, I-20133 Milano
[c]) Dipartimento di Chimica Strutturale e Stereochimica Inorganica, Università di Milano, I-20133 Milano

For *Albert* who will ever remain young enough to learn

Isomeric $C_{12}H_{12}$ hydrocarbon molecules with widely different constitution and shape are analysed for their packing ability. Some correspond to known compounds with known crystal structures, but some are invented hypothetical molecules designed to have low packing efficiency. For each isomer, a large number of close-packed, low-energy crystal structures was generated by computer, with lattice energies within a range of a few kJ mol^{-1}. Molecules with linear chains, triple bonds and Me groups tend to have larger molecular volume, lower lattice energy and lower crystal density than cyclic or cage isomers. The calculated crystal structures for each isomer show an inverse relationship between packing energy and cell volume. Although the slope dE/dV varies from molecule to molecule, the product of slope and free space stays roughly constant; less efficient crystal packings thus appear to be less sensitive to an increase in cell volume. Lattice-vibrational frequencies and the corresponding contributions to thermal vibrational entropy were estimated for real and virtual crystal structures. For a given isomer, as expected, a higher entropy goes with a larger cell volume, but different isomers show different entropy/volume relationships. At 300 K, $T\Delta S$ differences among computational polymorphs may compete with ΔH differences, thus making the lattice-vibrational entropy estimation a relevant factor in crystal-structure prediction.

Molecular Shape and Space Filling in Condensed Phases. – The boundary surfaces of atoms in molecules, regarded as solid bodies, are convex. Therefore, molecules cannot be packed so as to fill space without gaps or overlaps. Indeed, perfect space filling is possible only for a few simple geometric figures, such as parallelepipeds, triangular and hexagonal prisms, and rhombic dodecahedra. Since *Kepler*'s time it has been known that, in the closest packing of identical spheres, the ratio of occupied to total volume, the packing coefficient, is 0.74, and, since *Kitaigorodskii*'s pioneering work on molecular packing in crystals [1][2], we know that the close-packing principle applies also to organic crystal structures. Molecules in crystals tend to be surrounded by twelve to fourteen neighbouring molecules, the same as in typical close packed metals, and packing coefficients in molecular crystals vary only within a relatively narrow range, from *ca.* 0.65 to 0.80, not too different from the value for identical spheres. With a packing coefficient below *ca.* 0.6, substances are mostly in the liquid state, and below *ca.* 0.5 the attractive forces are no longer strong enough to hold the molecules together in a condensed state – the substance vaporises. Nevertheless, in spite of these regularities, organic molecules have very different shapes, and only a handful can be described as being even approximately spherical. It seems remarkable that molecules, whatever their shape, manage to pack with roughly the same efficiency as spheres.

Space Filling and Periodicity. – An understanding of crystal packing requires insight into the subtle relationships between intermolecular potential and crystal symmetry. To a first approximation, arguments about potentials can be replaced by arguments about shapes. Some purely geometrical points may therefore have a relevant chemical meaning. The conjecture that optimal space filling of a collection of identical (or enantiomorphic) objects of arbitrary shape necessarily leads to periodicity [3] has never been proved. It is related to *Hilbert*'s 18th problem, posed almost exactly 100 years ago [4]: '*How can one arrange most densely in space an infinite number of equal solids of given form, e.g., spheres with given radii or regular tetrahedra with given edges, that is, how can one so fit them together that the ratio of the filled to the unfilled space may be as great as possible?*' It is not clear from this whether *Hilbert* was referring to solids of arbitrary shape or only to regular solids, such as spheres or polyhedra. *Kepler*'s conjecture about cubic close packing of spheres can then be regarded as a special case.

Although it is fairly easy to prepare two-dimensional figures that tile the plane with space occupancy of 100%, this is possible only if symmetry is imposed [5–7]: non-periodic tilings such as *Penrose* tilings use at least two types of figure. There are no known quasi-crystals composed of single elements, only of alloys.

Molecular Shape and Crystal Packing: Isomers and Polymorphs. – The search for relationships between molecular shape and crystal packing has always been a challenging field of chemical and crystallographic research. Some simplification may be introduced when comparisons are restricted to crystal structures and properties among groups of isomers. Along these lines, a statistical study of packing variations among crystalline isomers has recently been made [8]. Crystals containing the same molecule in different packing arrangements are classed as polymorphs and have been studied statistically [9]. We can consider crystals of isomers as generalised polymorphs – the same atoms, combined in different bonding patterns *and* in different packing arrangements. Since we already have the concept of conformational polymorphism [10][11], the extension to constitutional polymorphism does not seem too far-fetched. The wider definition may not be to everyone's taste, but it serves our present purpose.

Indeed, there seems to be no generally accepted definition of polymorphism. The usual rule is that polymorphs are different crystal structures containing the same molecule. But what is the 'same' molecule? It has been taken to encompass the same molecule in different conformations, or even in different configurations (*e.g.*, α- and β-mannose, a well-known example of a disappearing polymorph [12]). And what about crystals containing different tautomers? Or the neutral and zwitterionic forms of an amino acid? The usual way of defining the borderline is to invoke interconvertibility in solution or in the melt [13], fast interconversion being associated with polymorphs, slow with different compounds. Presumably, this criterion is supposed to refer to room temperature. Thus, we do not usually regard crystals of *ortho*-, *meta*-, and *para*-dichlorobenzene as polymorphs, although they would be encompassed in our wider classification.

Crystal-structure determinations of polymorphs are scarce and haphazard, often carried out in different laboratories and under different conditions that are not always fully documented. From thermodynamics we know that among ordinary polymorphs, the low-temperature stable form must have the greater (more stabilising) packing

energy (lattice energy) and the smaller entropy, and it usually has the higher density. However, we are now in a position to study some aspects of polymorphism theoretically, by generating a large number (running into hundreds) of possible crystal structures for a given molecule, with the sole restriction that the structures correspond to minima in the potential-energy hypersurface, and studying their (calculated) properties. From such studies it appears that for a given molecule there are usually many packing arrangements with calculated packing energies within a range of a few kT, and that packing energy correlates broadly with crystal density. The root of this rule is that, as far as crystal-packing energy is concerned, empty space is wasted space; for calculated crystal structures of benzene [14], *e.g.*, the lattice energy decreases by *ca.* 0.8 kJ mol^{-1} for each additional $Å^3$ of free space; for a substituted coumarin [15] the equivalent figure is 0.6 kJ mol^{-1} $Å^{-3}$.

However, these calculated packing energies refer to atoms at rest. To allow vibrations of the molecules about their equilibrium positions and orientations in the crystal (and also for intramolecular vibrations) to occur, free space is required. At least for ordinary polymorphs, therefore, we can expect that intermolecular vibrational entropy should correlate with available free volume, that is, reciprocally with density. At very low temperatures, the entropic contribution to the free energy is negligible, but most organic crystal structures are observed at temperatures high enough that lattice-energy differences between polymorphs can be compensated by differences in entropy. There is thus a delicate balance between the two opposing effects of density on the thermodynamic stability of crystals. At the thermodynamic phase-transition temperature between polymorphs, the two are exactly equal. It must also be remembered that while our 'empty space is wasted space' motto refers mainly to dispersion or *van der Waals* energy contributions, other intermolecular interactions may be brought into play, notably the H-bond. Because of its strong directional preference, this may tend to produce open rather than tight packing arrangements. Water is the obvious example, but one that can hardly be generalised – water is such an extraordinary, actually unique substance. Other possibly relevant influences may be those involving weaker or 'non-conventional' H-bonds, or π-π interactions, dipole-dipole and higher multipole interactions, and a host of others, although no trace of dependence of average packing coefficients on the nature of such intermolecular interactions in crystals has ever been firmly detected [16].

Scope of this Paper. – Among crystal properties of interest here, energy, entropy and density substantially depend on molecular mass. We are interested here in shape effects, *i.e.*, the effects of different geometrical deployments of the same mass. Thus, we are led to examine and compare crystal structures of isomers, of which conformational polymorphs are a subclass. We chose the $C_{12}H_{12}$ groups because of the relatively large density differences found among crystal structures of its members, and because the molecular size lends itself nicely to extensive computations, but other groups might have served as well. For the experimentally determined structures, we estimate quantities such as packing coefficient, lattice energy and rigid-body vibrational entropy, and try to relate these crystal properties to molecular shape. In addition, for each of the isomers we generate large numbers of alternative crystal structures and compare their properties among themselves and also with the experimental structures. Finally, we do

not restrict ourselves to compounds for which experimental data are available; we designed a few hypothetical $C_{12}H_{12}$ molecules to have maximally awkward shapes for packing, and calculated a number of possible crystal structures.

A program such as this is only possible because computational and simulation methods are now available for rapidly generating a diversity of possible crystal structures from a molecular diagram and for estimating their lattice energies and rigid-body lattice-vibrational entropies. We can now speak about 'virtual crystallography', where hypothetical crystal structures and properties are based exclusively on computer calculations, and also about 'semi-virtual crystallography', where calculations of crystal properties are based on experimentally determined crystal structures. In this paper, we seek a synergy of these approaches to the problem of correlating, in a general way, molecular shape or at least some molecular geometrical parameters, with lattice energy, lattice-vibrational entropy and crystal density.

Methods. – A recent survey of crystal densities in several isomer groups [8], retrieved from the *Cambridge Structural Database* (*CSD*)[1]), revealed striking density variations of up to 20% in the $C_{12}H_{12}$ group, from 1.119 g/ml to 1.341 g/ml. For each molecule and crystal, we calculate the crystal density, $D_c = M_{mol} \cdot Z/V_{cell}$, where M_{mol} is the molecular mass (g) and Z is the number of molecules in the crystal unit cell, together with the molecular volume, V_{mol} (Å^3), defined as the volume of a single molecule. In crystallographic contexts, molecular volume is sometimes defined as the unit cell volume divided by the number of molecules in the cell, V_{cell}/Z; the two volumes are related by the *Kitaigorodskii* packing coefficient $C_k = Z \cdot V_{mol}/V_{cell}$.

Since molecules have no strictly defined boundaries, the molecular volume is not a rigorously defined quantity but depends on somewhat arbitrary choices of atomic radii. Here it is estimated by standard methods [17][18]. It appears that, no matter what method is used, large/small relationships hold. Other useful indicators of molecular compactness are the molecular density, $D_{mol} = M_{mol}/V_{mol}$, and N_B, the number of linkages (bonds of any multiplicity) between non-H-atoms in the molecule [8].

When atomic coordinates and space-group symmetry are known, the packing energy E (in kJ mol^{-1}) is calculated by pairwise additive atom-atom potentials, *i.e.*, $E = \Sigma_i \Sigma_j A \exp(-BR_{ij}) - C R_{ij}^{-6}$, where i runs over atoms of a reference molecule, and j over atoms of all surrounding molecules. Parameters A, B, C have been calibrated for organic crystals [19]. Note that E is always stabilising, *i.e.*, $E < 0$, but for simplicity it is reported here as positive. Lattice sums were cut at 10 Å. Experience has shown that this threshold yields *ca.* 97% of the total convergent energy.

With the same potentials, the lattice-dynamical (intermolecular) contribution to the crystal vibrational entropy, S_{ext}, can also be estimated in the rigid-body approximation [20]. The calculations provide the first and second derivatives of the potential energy, *i.e.*, the forces and force constants, with respect to the six degrees of freedom

---

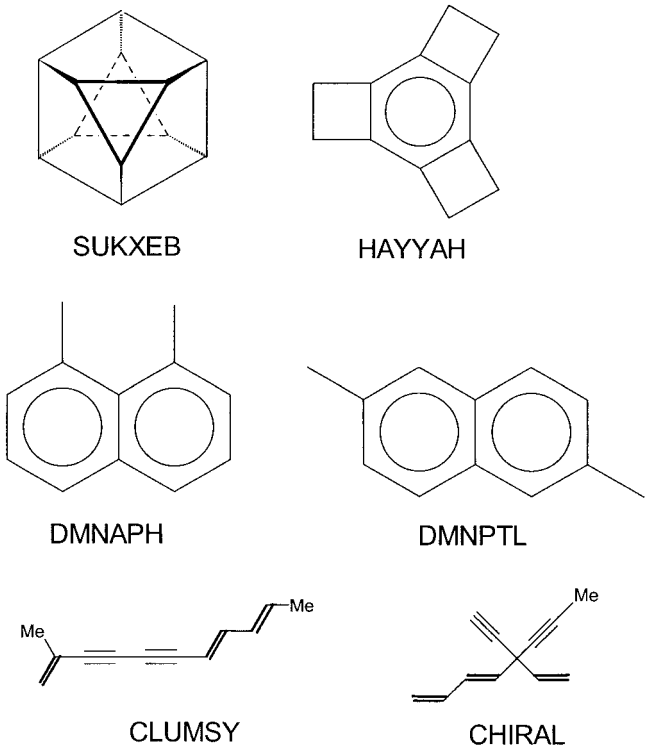
[1]) The *Cambridge Crystallographic Data Centre*, *CCDC*, 12 Union Road, Cambridge CB2 1EZ, England, produces and distributes the *Cambridge Structural Database* (*CSD*), which contains, in computer readable form, unit-cell dimensions, space groups, atomic coordinates and bibliographic data for organic and organometallic crystals of known structure, plus software for retrieval, visualisation and statistical analysis of the results. The *CSD* contains more than 200,000 entries and is expanding at *ca.* 20,000 entries annually.

corresponding to displacement of the molecular centre of mass and libration around the inertial axes. *Newton*'s equations of motion are then solved in the harmonic approximation to yield, for the equilibrium structure, mean-square displacements, libration amplitudes and corresponding frequencies. The procedure does not differ much from the one applied in molecular spectroscopy, except that lattice periodicity has to be taken into account. This is done by introducing a reciprocal space wave vector, **k**, and calculating the frequencies by sampling over the independent region of **k** space, the *Brillouin* zone. Apart from this technicality, the calculations leading to enthalpy, entropy and heat capacity are made with the usual formulas of statistical thermodynamics, based on the vibrational partition function. It must be stressed that the calculated quantities refer only to the rigid-body, intermolecular part of the total crystal vibration, and are therefore sometimes called 'external' thermodynamic functions. The intramolecular part is roughly constant in different crystal structures of the same molecule, at least for the semi-rigid molecules we consider here.

The above calculations are examples of what could be called semi-virtual crystallography; since they are based on known crystal structures, they provide reasonably sound information for a small investment in computer time. Turning now to virtual crystallography, computer generation of possible crystal structures shows that, for a given molecule, there are often many crystal arrangements with almost the same packing energy. That is a reason why purely *ab initio* crystal-structure prediction is so difficult [21]. On the other hand, the possibility of generating a large number of hypothetical crystal structures for the same molecule, and the comparison of their densities and energies, provides a good opportunity to study the hows and whys of crystal packing.

For each isomer considered here, a molecular structure was assumed as taken from the crystal-structure determination, or, for the hypothetical isomers CHIRAL and CLUMSY (*Scheme*), using standard bond lengths (triple bond 1.19 Å, double bond 1.31 Å, single bonds: between triple bonds, 1.38 Å, between double bonds, 1.46 Å, between a triple and a double bond 1.43 Å, adjacent to triple bond 1.47 Å, adjacent to double bond or C-methyl, 1.50 Å, angles of 180, 120 and 109.47° as applicable, all C–H distances 1.08 Å). A diversity of possible crystal structures was then generated, using a 'polymorph predictor' package, *Zip-Promet* [22]. Briefly, the procedure consists in building dimers, ribbons or layers of molecules according to partial space group symmetry operations, selecting the most cohesive among such substructures, and applying further translations as required to obtain complete, but still approximately packed, three-dimensional crystal structures. Only the most popular space groups for organic crystals were considered, *i.e.*, $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12_1$, *Pbca*, $C2/c$, and only structures with one molecule or less in the asymmetric crystal unit. Finally, each of these raw structures was optimised with respect to variation of cell parameters and rigid-body molecular degrees of freedom using a recently developed algorithm [23]. These calculated structures are 'temperature-less', in the sense that no temperature is ever specified in the whole computational procedure, and, consequently, no allowance for molecular motion is included. For comparisons with experimental crystal structures, the latter have to be 'optimised' by exactly the same procedure. A unit-cell shrinkage of a few percent invariably results as the experimental structure is formally brought to zero temperature; such optimised structures are labelled 'OPT'. In the real world, many

Scheme



room-temperature crystal structures undergo phase transition on cooling, often to closely similar structures showing lower symmetry, *i.e.*, with a larger asymmetric unit. Indeed, at least one of the isomers considered here, HAYYAH, shows such behaviour. Although the driving force for these symmetry-breaking structural changes on cooling is obviously towards better packing energy and hence, very likely, also towards higher packing density, no allowance for this is made in our calculations. For each crystal structure considered here, whether computer-generated or experimental, the lattice-vibrational entropy was calculated as outlined above. The entropy thus calculated is a hybrid quantity, in the sense that although it refers to the temperature (usually 300 K) used in the statistical mechanics calculation, it is based on a 'temperature-less' crystal structure with no allowance for thermal expansion.

It is sometimes useful to have an estimate of the density of a virtual crystal structure at room temperature. Thermal expansion coefficients were calculated from unit-cell data for a sample of *ca.* 50 variable-temperature organic crystal structures listed in the *CSD*. Values of $\alpha = 1/V\,(\mathrm{d}V/\mathrm{d}T)$ cluster around $2 \cdot 10^{-4}\,\mathrm{K}^{-1}$ with a few outliers and some scatter. Using this estimate for our $C_{12}H_{12}$ crystals, warming from 0 to 300 K corresponds to a *ca.* 6% increase in cell volume, or, equivalently, to a 6% decrease in density and packing coefficient.

Results: Energies and Cell Volumes. – The *Scheme* shows the $C_{12}H_{12}$ isomers that were considered: throughout the paper, we refer to known compounds using their six-letter *CSD* label (refcode), and to unknown ones using refcodes invented by us. Their molecular volumes and densities (*Table 1*) cover a range of more than 20% from compact hexacyclic SUKXEB to flat polycyclic HAYYAH to the flat aromatic dimethylnaphthalenes to the hypothetical isomers CLUMSY and CHIRAL, designed to have especially awkward shapes for packing. Molecular density increases with N_B, the number of C,C linkages [8].

Table 1. *Some Properties of the $C_{12}H_{12}$ Isomer Molecules, and Global Results of the Crystal-Structure Generation Procedures*[a])

Molecule	D_{mol}	V_{mol}	N_B	No. of structures generated	Slope dE/dV
SUKXEB	1.776	145.9	18	137	– 0.52
HAYYAH	1.727	158.1	15	78	– 0.50
DMNAPH	1.606	161.3	13	164	– 0.40
DMNPTL	1.596	162.3	13	365	– 0.45
CLUMSY	1.469	173.2	11	43	– 0.37
CLUMSY90	1.469	173.2	11	139	– 0.34
CHIRAL	1.459	177.6	11	156	– 0.38

[a]) Energies in kJ mol^{-1}, volumes in Å^3, density in g ml^{-1}. See text for definitions of symbols.

Results of crystal-structure generation runs are best visualised in the form of plots for the lattice energy, E, or vibrational entropy, S_{ext}, against the normalized cell volume, V_{cell}/Z, for a large number of calculated crystal structures in several space groups. *Table 2* summarises numerical results for the most stable crystal structures observed and calculated for each isomer.

Figs. 1–3 show the energy-entropy-volume relationship for the three compounds with top, middle and lowest molecular density in *Table 1*. There is always a clear inverse relationship between energy and cell volume, showing that energies become more stabilising as molecules come closer together. This relationship is typical and holds for all isomers studied, independent of space group, at least within the pool of close-packed ones here considered. The structure-generator algorithm ensures that each calculated structure is a minimum in the lattice potential-energy hypersurface, and it is remarkable that such minima are present at densities 15% lower than that of the most stable structure[2]). However, given that the volume increase accompanying melting is typically in the 10–15% range for organic compounds, this result is not

[2]) The computer builds polymorphs without regard to actual paths that interconnect them through the energy hypersurface. These structures are stable towards transformation because of the high activation energy required to disentangle the molecules from the structural interlock. Paths towards crystal transformation require the overcoming of both attractive and repulsive forces, those involved in striking the zero-force balance attained at equilibrium.

Table 2. *Lattice Energies, Normalised Cell Volumes, Packing Coefficients and Densities of the Experimental and of the Most Stable Computer-Generated Crystal Structures for the Molecules in the* Scheme

	E [kJ mol^{-1}]	V_{cell}/Z [Å^3]	C_k (C_k estimated at 300 K)	Density [g/ml] (relative %)
SUKXEB exp, OPT[a])	**83.6**	**187.4**	**0.779 (0.735)**	**1.382 (0)**
$P\bar{1}$	80.7	191.5	0.762	1.353
$P2_1/c$	79.1	193.5	0.754	1.339
$P2_1$	80.7	190.7	0.765	1.358
$P2_12_12_1$	79.5	192.7	0.757	1.344
HAYYAH $P2_1/c$ (= exp, OPT)[b])	**83.7**	**208.6**	**0.758 (0.715)**	**1.242 (– 10%)**
$P\bar{1}$	82.5	211.4	0.748	1.225
DMNAPH $P2_1/c$	**83.5**	**215.0**	**0.750 (0.707)**	**1.205 (– 13%)**
$P2_1/c$, exp, OPT[c])	81.1	217.0	0.743	1.194
$P\bar{1}$	82.0	217.3	0.742	1.192
$P2_12_12_1$	81.5	217.7	0.741	1.190
DMNPTL *Pbca* (= exp, OPT)[d])	**84.5**	**220.9**	**0.735 (0.693)**	**1.173 (– 15%)**
$P\bar{1}$	83.0	219.0	0.741	1.183
$P2_1/c$	84.1	216.7	0.749	1.195
$P2_1$	84.4	221.5	0.732	1.169
CLUMSY $P2_1/c$	**80.2**	**247.4**	**0.700 (0.660)**	**1.047 (– 24%)**
	80.0	250.0	0.693	1.036
CLUMSY90 $P\bar{1}$	**77.8**	**251.9**	**0.687 (0.648)**	**1.028 (– 26%)**
$P2_1/c$	78.7	252.3	0.686	1.027
CHIRAL $P\bar{1}$	**76.4**	**243.2**	**0.730 (0.688)**	**1.065 (– 23%)**
$P2_1/c$	74.4	245.4	0.724	1.056
$P2_1$	66.5	258.0	0.683	1.005
$P2_12_12_1$	68.0	252.5	0.698	1.026

[a]) From [25]. [b]) From [29]. [c]) From [31]. [d]) Half molecule in the asymmetric unit [30]. All other results refer to one molecule in the asymmetric unit.

unreasonable and shows that the purely empirical potentials used here adhere to the implied physics.

The similar plots obtained for all isomers (see also *Figs. 4–6*) demonstrate that many structural possibilities exist for crystal aggregation of a given molecule, as might be expected for the packing of bodies of arbitrary shape and for interaction potentials with little directional selectivity. The result is, however, general and independent of the type of molecule and choice of force field. Indeed, quite similar results have been obtained for structures with the strongest directional pointer in organic chemistry, the H-bond, as in crystals of monosaccharides [24]. A phase space with many energy minima, separated by high energy barriers, seems to be an inherent feature of the crystal packing of organic compounds. This may help to explain why polymorphism is so common among organic compounds, why first-order solid-solid transitions are so sluggish, and why these hardly ever transform in a single-crystal to single-crystal manner.

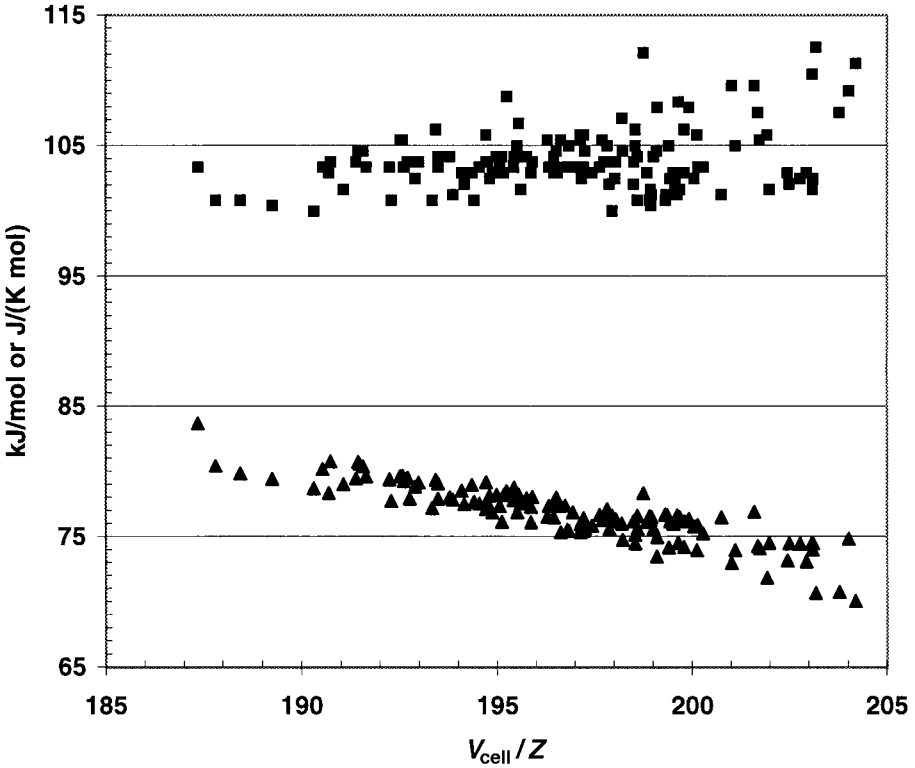


Fig. 1. *Plot of lattice energy*, E (triangles), *and lattice-vibrational entropy*, S_{ext} (squares), vs. V_{cell}/Z *for the virtual, computer-generated crystal structures of isomer SUKXEB* (*Scheme*). The isolated point at the upper left of the distributions corresponds to the observed crystal structure (exp, OPT in *Table 1*)

For the highly compact molecule SUKXEB, our search procedure generated 13 crystal structures in four different space groups within a 2 kJ mol^{-1} packing-energy range and a 3% density-difference range. Here, there is no question of a correct or incorrect prediction because the space group of the experimental crystal structure [25], $R\bar{3}c$ (No. 167), is not one of those considered in the generation procedure. None of the monoclinic or orthorhombic structures generated in our search gets close to the stability of the experimental crystal structure, a negative result that nonetheless confirms the reliability of our potentials. The presence of a threefold axis in the crystal packing seems to be necessary for optimal packing of this threefold symmetric molecule.

For contrast, we turn to the crystal structures generated for some of the compounds with supposedly poor packing abilities. The worst packer of all, according to our criteria [8], should be CHIRAL (*Scheme*) because of the absence of cycles and the presence of triple bonds, which lower the value of parameter N_B, and because of the branched structure and the presence of Me groups. Indeed, as seen from *Table 2*, the increase in cell volume amounts to 30%. The increase in molecular volume V_{mol} is only 21%, so not

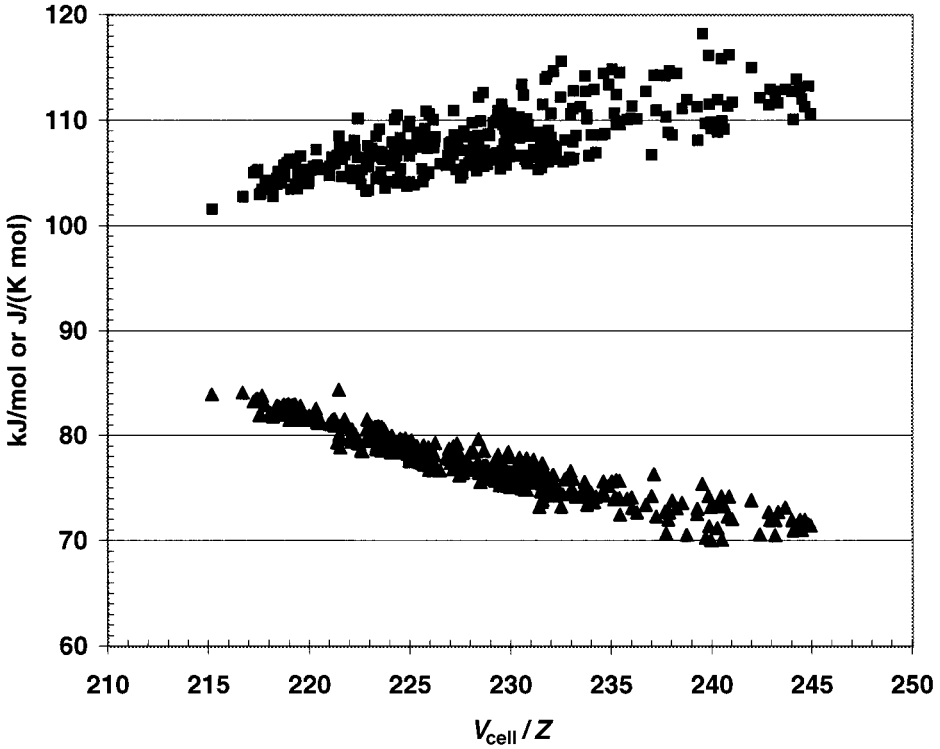


Fig. 2. *Plots of lattice energy*, E, *and lattice-vibrational entropy*, S_{ext} (as in *Fig. 1) vs.* V_{cell}/Z *for the virtual crystal structures of isomer DMNPTL* (*Scheme*)

only does CHIRAL occupy more space than SUKXEB, it also packs less efficiently due to its more awkward shape. This is also expressed by the decrease of the packing coefficient from 0.78 to 0.73. The packing energy of the best CHIRAL structure is 76.4 kJ mol^{-1}, compared with 83.6 kJ mol^{-1} for SUKXEB. The most stable structures generated for CHIRAL occur in centrosymmetric space groups and thus require the presence of both enantiomers. Structures in the chiral space groups $P2_1$ and $P2_12_12_1$ are much less stable (see *Table 2*), indicating that the centre of symmetry is the best packing operator for molecules with awkward shapes. Presumably, the centrosymmetric dimer offers a less awkward shape than the monomer.

For CLUMSY, two conformations were considered, one planar and one with a 90° rotation around a C–C bond (see *Scheme*). Only two space groups were examined, $P\bar{1}$ and $P2_1/c$. *Fig. 5* shows that the crystal structures for the 90° conformer are generally less dense and less stable than those for the planar one, suggesting that the 90° conformer packs less efficiently. Since the molecular volumes are the same for these two conformers, this is a pure shape effect on the crystal-packing energy. Note that this is an extreme case of virtual structure, since rotation around the C–C bond is essentially free, and nothing could prevent the molecule from assuming a planar shape at crystallisation time.

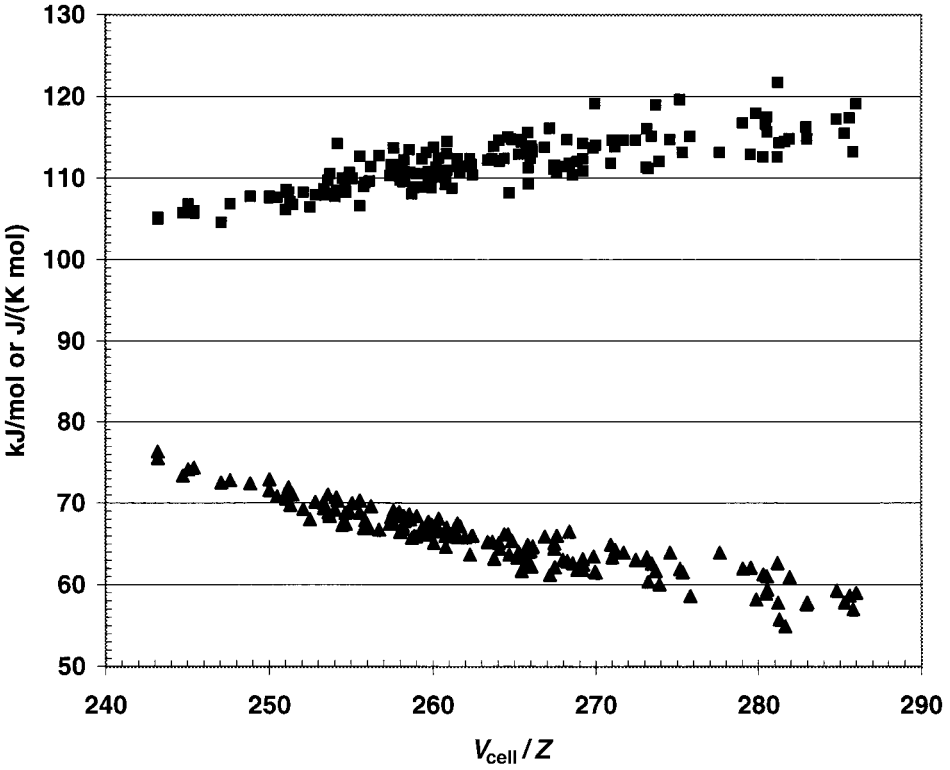


Fig. 3. *Plots of lattice energy*, E, *and lattice-vibrational entropy*, S_{ext} (as in *Fig. 1*) vs. V_{cell}/Z *for the virtual crystal structures of isomer CHIRAL (Scheme*)

Fig. 6 shows the corresponding plots for the isomeric dimethylnaphthalenes. As previously noted [26], the centrosymmetric 2,6-isomer packs more favourably, calculated structures at any cell volume being slightly more stable than those for the 1,8-isomer. Again, this is a pure shape effect in crystal packing, due here to constitution rather than to conformation.

Molecular Shape and Packing Fitness. – *Tables 1* and *2* show how the isomers rank with respect to their overall packing ability. In crystal density, tetracyclic HAYYAH with its aromatic core ranks second after SUKXEB, followed by DMNAPH and DMNPTL with their Me groups, which are detrimental to packing efficiency. The most impressive result of these crystal-structure generation exercises is perhaps the very large (26%) decrease in crystal density on going from SUKXEB to CLUMSY. This comes from a combination of two factors, a 20% increase in molecular volume (from 146 to 173 Å^3) and a 10% decrease in packing coefficient (from 0.78 to 0.71). Despite this large spread in crystal densities, all generated structures have packing coefficients close to 0.74, demonstrating the ability of almost any molecular shape to find its way to a close-packed structure. Although the 10% range in packing coefficient may seem small from one point of view, it covers almost the whole range of packing coefficients

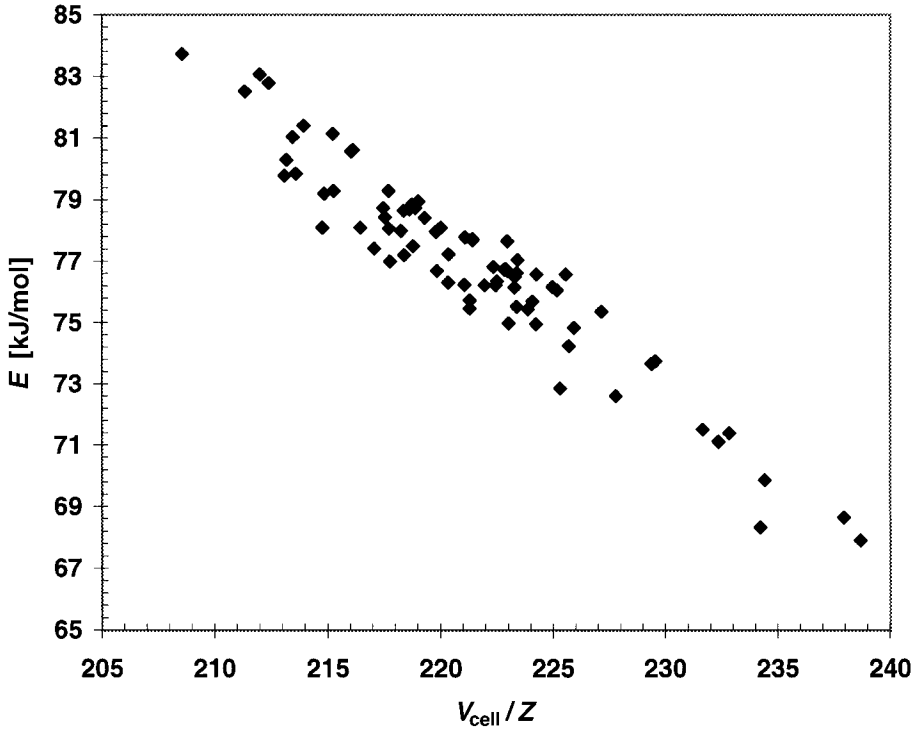


Fig. 4. *Lattice energy*, E, vs. V_{cell}/Z *for the virtual crystal structures of isomer HAYYAH* (*Scheme*)

observed in organic compounds. In fact, if the packing coefficients are corrected for the estimated 6% increase in cell volume on going from 0 to 300 K, the range is 0.660–0.735, just about the extremes of the distribution obtained for general organic molecules [16]. As noted above, isomers with a smaller molecular volume are inclined to pack more tightly; density and packing coefficient follow the trend that relates an increase in packing fitness with an increase in D_{mol}.

Although SUKXEB has by far the highest density, the packing energy of its optimised experimental structure is no better than for HAYYAH, DMNAPH or DMNPTL, and the packing energies of its virtual structures are worse and comparable even with those of CLUMSY. This appears to go against the general tendency evident in the various *E*/*V* plots that packing energy increases with density, and therefore calls for explanation. Since the C-atoms in SUKXEB are saturated, while those in the other molecules mentioned are not, one might think this difference is relevant, but pair potentials in our force field make no distinction betwen aliphatic, unsaturated and aromatic C-atoms, an approximation which has been found excellent so far. The main contribution to the packing energies comes from C···H interactions. The H···H interactions contribute very little, and, in the SUKXEB crystal structures, the short intermolecular contacts are mainly of this type. The methine C-atoms in this molecule are shielded from optimal intermolecular contacts by their attached H-atoms. On the other hand, six of the twelve C-atoms in HAYYAH, five in CLUMSY, and four in

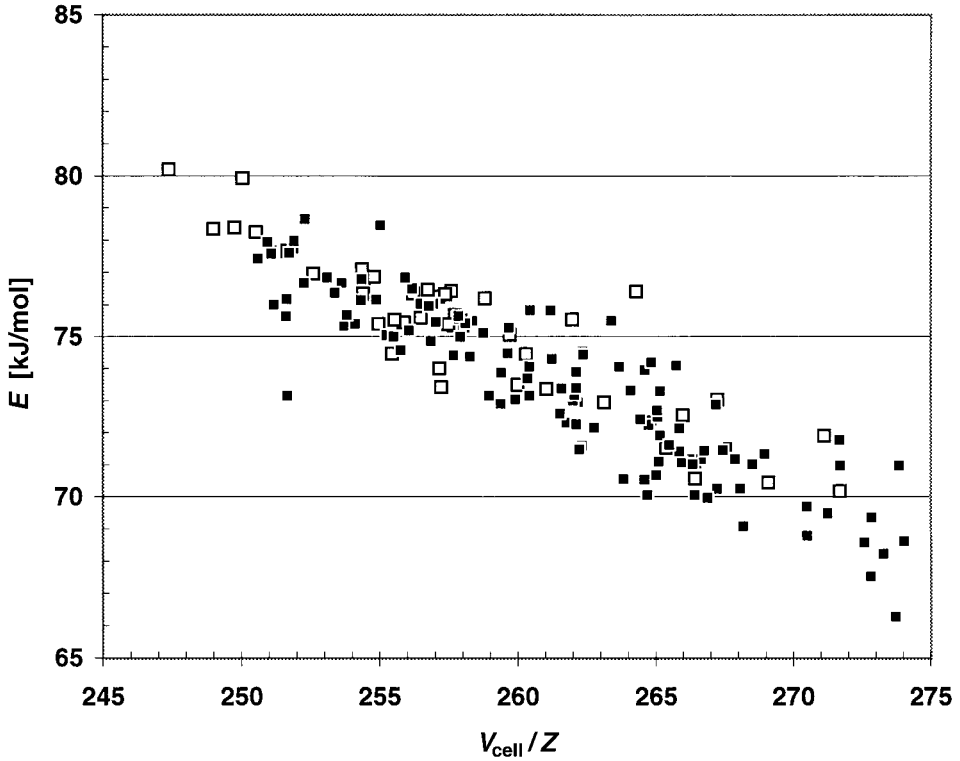


Fig. 5. *Comparative plots of lattice energy*, E, vs. V_{cell}/Z *for the virtual crystal structures of isomers CLUMSY* (open squares) *and CLUMSY90* (black squares)

DMNAPH and DMNPTL have no attached H-atoms and are not hindered from favourable intermolecular contacts in this way.

Another molecular-shape factor plays a role as far as the intermolecular contacts are concerned. In SUKXEB each C-atom can enter into close stabilising intermolecular contacts only in directions that point outward from the molecular surface; other directions are shielded by the molecule itself. In planar aromatic HAYYAH, DMNAPH and DMNPTL the C-atoms can make such contacts in three directions, outwards and on both sides of the molecular plane. In CLUMSY the C-atoms can achieve intermolecular contacts in four directions, so to say, up and down, east and west. In CLUMSY90 there is only a small region where one can speak of an inside and an outside, and only the outside contributes to the packing energy. CHIRAL pays a penalty by having its central quaternary atom completely, and the adjacent C-atoms partialy shielded from the outside world.

The above discussion shows that the influence of molecular shape on packing density and on packing energy do not necessarily run parallel. A compact molecule gives the highest packing density but not necessarily the highest packing energy. We have previously noted [3] that the sublimation enthalpy of C_{60} is just half of that expected for a planar aromatic compound with the same number of C-atoms. Among

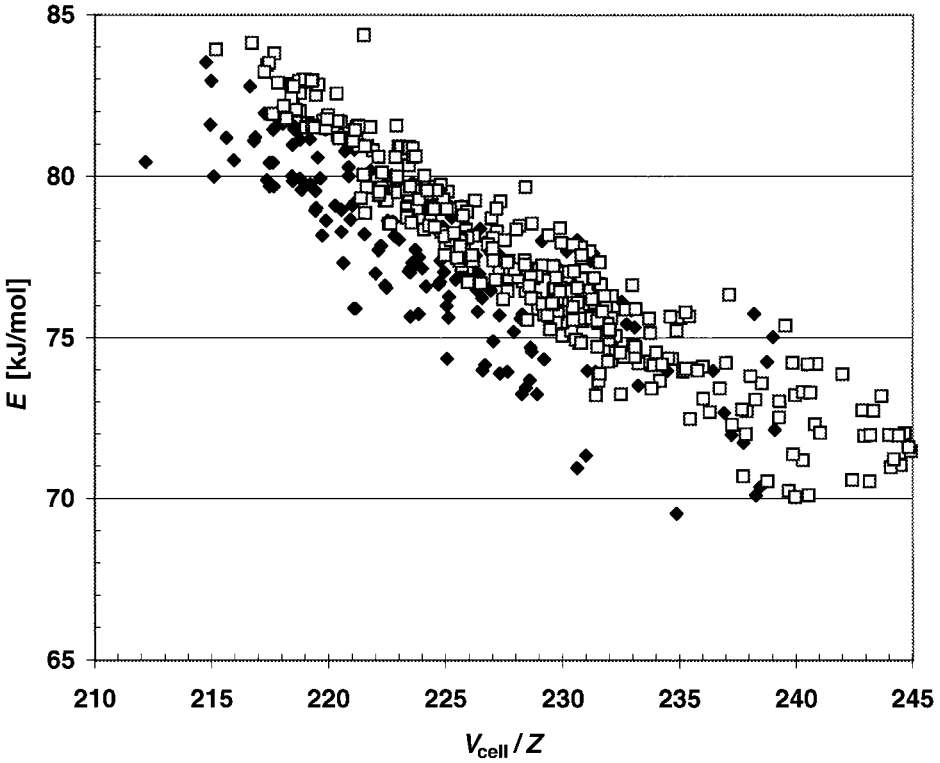


Fig. 6. *Comparative plots of lattice energy*, E, vs. V_{cell}/Z *for the virtual crystal structures of isomers DMNAPH* (black squares) *and DMNPTL* (open squares)

polymorphs (different packings of the same molecular shape) density and lattice energy run parallel, but among isomers (different molecular shape) this is not necessarily the case.

Absolute Close-Packing. – The approximately linear appearance of the E/V plots suggests an extrapolation and re-interpretation of the close-packing principle. One may ask, for example, what would be the energy of a hypothetical crystal structure with a packing coefficient of unity, that is, a structure in which the occupied volume is equal to the molecular volume so that there is no free space – a sort of three-dimensional *Escher* pattern? We ignore the physical impossibility of deforming the molecules into the geometric shapes that allow such perfect space filling. We can rewrite the E/V relationship

$$E = (\mathrm{d}E/\mathrm{d}V)V/Z + E_0$$

in terms of $V_{\mathrm{free}} = (V/Z) - V_{\mathrm{mol}}$ and estimate the value of the energy for which $V_{\mathrm{free}} = 0$. With the information given in *Tables 1* and *2* for the slopes $(\mathrm{d}E/\mathrm{d}V)$, intercepts $E(0)$, and molecular volumes V_{mol} for the various isomers, the estimated

energies for these hypothetical perfect close packed structures all lie within the range 100–113 kJ mol^{-1}, *ca.* 20% greater than the values obtained for the best calculated structures. It would be an interesting challenge to design the $C_{12}H_{12}$ isomer that most nearly approaches this limit or, at least, an isomer that can pack with a higher packing coefficient than SUKXEB, the present record holder with $C_k = 0.78$. Of course, for actual crystals made from actual molecules, compression is not accompanied by a gain in energy but the reverse.

In the above exercise, the slope (dE/dV) varies from one isomer to another, over a range of 0.34 to 0.52 kJ mol^{-1} Å^{-3}. The slope is a kind of sensitivity indicator, describing how much the various structures calculated for each isomer lose in energy as the cell volume increases. We find that the slope varies inversely with the amount of free space $V_{free} = (V/Z) - V_{mol}$ for the densest crystal structure obtained or calculated for each $C_{12}H_{12}$ isomer. The data in *Tables 1* and *2* show that the product $(dE/dV)V_{free}$ lies between 22 and 29 kJ mol^{-1} for the various isomers, that is, it is roughly constant. If we take it as constant, then by extrapolation, as V_{free} tends to zero, (dE/dV) tends to infinity. This implies that if we could find a crystal structure with the density of a perfectly close packed structure, then any alternative structure, even one showing only a slight increase in cell volume, would have a much lower lattice energy. The perfect space-filling arrangement would then be unique, not only geometrically but also from a packing-energy viewpoint. As the amount of free space in the 'best' crystal structure increases, so does the number of alternative structures with only slightly less lattice energy. For molecules that can only find packing arrangements with much larger amounts of free space, the sensitivity indicator (dE/dV) becomes very small, up to a point where it become negligible and the structure behaves like a liquid.

Lattice-Vibrational Entropies. – For different packing arrangements of a given molecule, increase in cell volume can be expected to correlate with a loss of packing energy and an increase in vibrational entropy. A less tightly packed crystal structure should favour lower-frequency lattice vibrations and hence lead to an increase in the thermal (energy dispersion) contribution to the entropy. This is in agreement with the general rule that, among polymorphs, the low-temperature form has the higher density [27]. Computer generation of many crystal structures for the same molecule is the ideal computational experiment to test these expectations. Moreover, the confrontation of calculated entropies and energies should provide clues concerning free energies and hence relative stabilities of postulated crystal structures, which may be relevant if the calculations are done with crystal-structure prediction in mind. However, since the calculated crystal structures are formally 0 K structures, such comparisons involve the assumption that entropy *differences* among competing crystal structures remain essentially constant with thermal expansion, an assumption that seems to be substantiated by actual temperature-dependent entropy calculations [28].

The plots shown in *Figs. 1–3* confirm and illustrate the expected relationships between volume, energy and entropy. For the CHIRAL structures, the *E vs. S* plot (*Fig. 7*) also shows the expected trend: a more stabilising energy goes with a smaller entropy. Entropy differences among the more stable crystal structures are small. Only when the lattice energy of the CHIRAL structures becomes less than *ca.* 70 kJ mol^{-1} does the scatter increase significantly; only as more and more packing options become

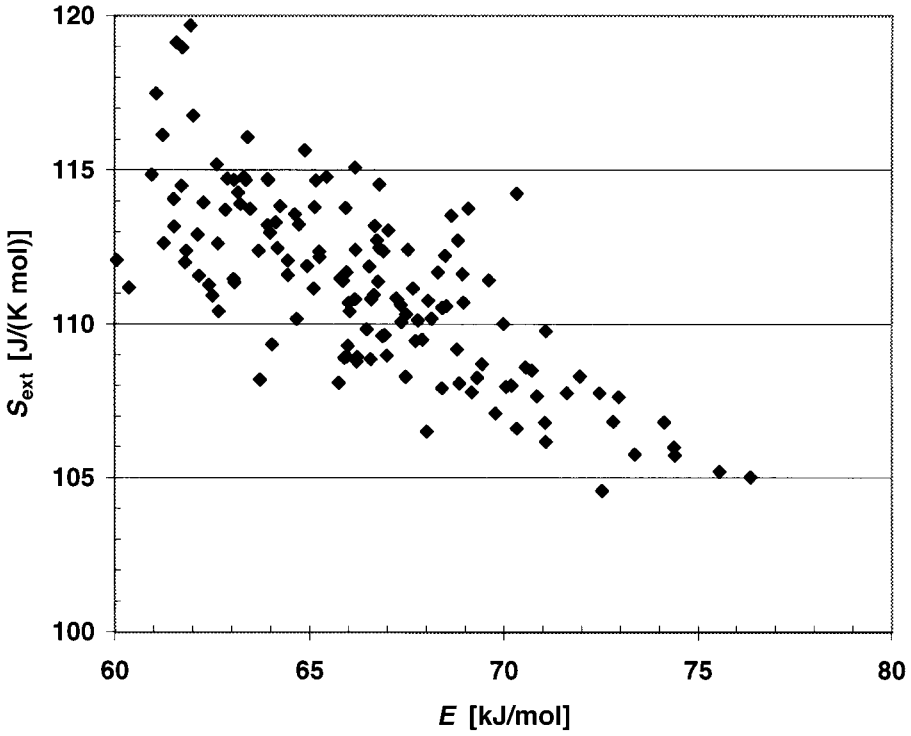


Fig. 7. *Lattice-vibrational entropy*, S_{ext}, vs. *lattice energy*, E, *for the virtual crystal structures of isomer CHIRAL* (*Scheme*)

available in a less restricted space do entropy contributions compete with enthalpy differences in the free energy at 300 K. For this compound, such considerations definitely belong to the virtual world and entropy differences are hardly relevant in crystal structure prediction.

Fig. 8 shows the *E vs. S* plot for SUKXEB, the isomer at the other extreme of the shape range. Here, the experimental crystal structure has by far the best lattice energy and a comparatively large entropy, so that it is stabilised on both counts. The scatterplots in *Fig. 8* and in the upper part of *Fig. 1* show, however, that, among virtual crystal structures, the *S*/*V* and *S*/*E* correlation is almost lost. This implies that structures with a low vibrational entropy can occur even with a less compact packing, or, *vice versa*, that very compact structures with a highly stabilising lattice energy can still be associated with a comparatively large vibrational entropy. There is no straightforward explanation for this behaviour, and any attempt to interpret vibrational entropies in terms of either molecular- or crystal-structural features would hardly go beyond speculation, without a quantitative analysis of vibrational eigenvectors. While this task is left for ongoing work, we take this result as an indication that at least for some molecules the entropy term can compete with enthalpy terms, and therefore should not be neglected *a priori* in computational crystal-structure prediction.

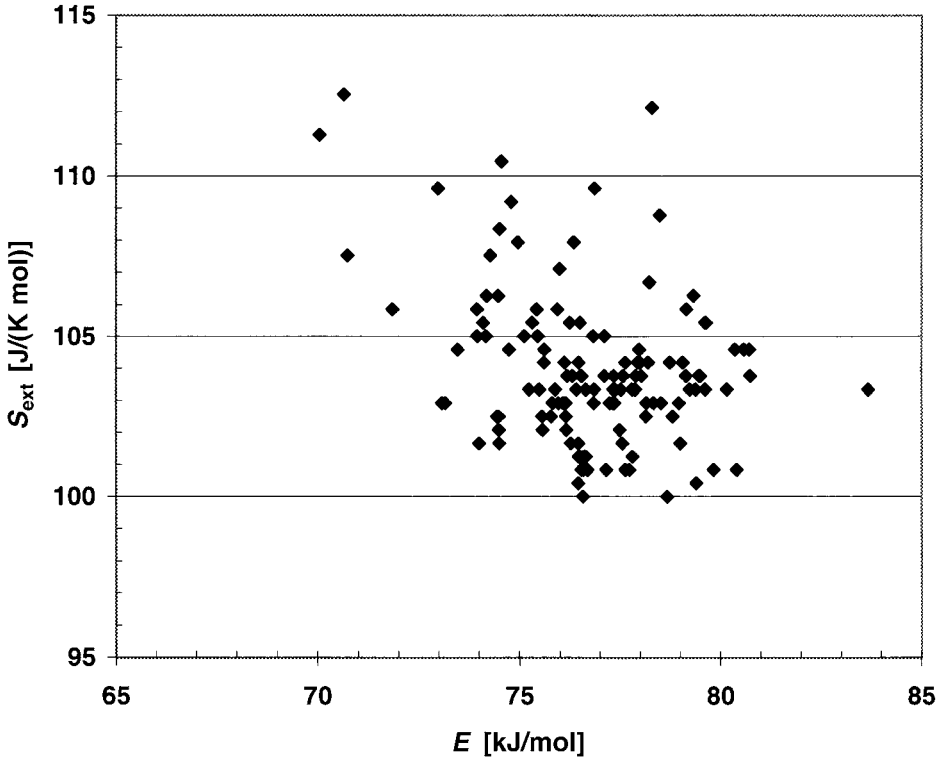


Fig. 8. *Lattice-vibrational entropy*, S_{ext}, vs. *lattice energy*, E, *for the virtual crystal structures of isomer SUKXEB* (*Scheme*)

Summary. – In our search for correlations between molecular properties and crystal structure, we have analysed a group of isomeric $C_{12}H_{12}$ hydrocarbons with widely different molecular constitution and shape. The group contains compounds with known crystal structures, but we have also included hypothetical isomers with constitutions specifically designed to bring about lower crystal density, according to previously derived rules. Many crystal structures were computer-generated for each of these isomers and their lattice energies, densities, packing coefficients and vibration entropies compared and contrasted.

Plots of lattice energy and vibrational entropy against cell volume show the expected trend: as more free space is available, and the packing is less tight, more structures become possible, while the lattice energy becomes less stabilising and the entropy increases. Our results confirm that, for a given molecular structure, many close-packed, low-energy crystal structures can be generated within a range of a few kJ mol^{-1}. However, the density of the best structures is very different among isomers, up to 30%. Polycyclic globular molecules form the most dense crystals, isomers with linear chains, triple bonds and Me groups the least dense. In particular, the results of the crystal-structure generation for the hypothetical molecules CLUMSY and CHIRAL show that planar molecules pack better than twisted ones, and that the centre of symmetry seems

to be an indispensable feature for optimal packing of chiral molecules of awkward shape. This kind of analysis using 'virtual crystallography' computer experiments confirms that our criteria for predicting the ease of packing on the basis of molecular shape are valid, at least in hydrocarbons. Packing fitness results primarily from shapes that occupy the smallest possible volume.

For each of the various isomers studied, there is an inverse relationship between lattice energy E and cell volume V_{cell}/Z. The product of the slope dE/dV and the free volume in the cell of the most stable crystal structure is roughly constant from one isomer to another. In other words, the more tightly packed structures have a higher slope, while more open crystal packings are relatively insensitive to further increase in the amount of free space. Extrapolation of the E/V relationship suggests that observed crystal structures can reach *ca.* 80% of the lattice energy expected for idealised perfect close packing, with complete space occupancy and a packing coefficient of unity. Ways of approaching this ideal within the rules and restrictions of chemistry have still to be found.

While density and lattice energy run parallel among polymorphs (real and virtual), this is not necessarily the case among isomers with different molecular shapes. There, crystal structures of greatly differing density can have nearly the same lattice energy. This shows that packing-efficiency arguments, based purely on molecular shape, cannot transfer directly into energetic arguments. The influence of molecular shape on crystal density is easier to understand than its influence on packing energy.

Our study in computer generation of crystal structures of the various isomers illustrates the whole range of cases encountered when trying to predict crystal structures. For SUKXEB, any attempt at prediction would have been doomed to failure, because the space group of the experimental structure is not one of the most frequent in organic crystals and is not considered in our procedure. For HAYYAH and DMNPTL, the calculated structure with the lowest lattice energy corresponds to the experimental crystal structure ([29] and [30], resp.), so that these examples would rank as successful predictions, according to the present rules of the game, which do not consider kinetic or entropic effects. For DMNAPH, the experimental crystal structure [31] was nowhere among the calculated ones, a complete prediction failure. The reasons for these differences in predictive power remain obscure. We may claim a modest success rate in this rather limited sample, and leave our postulated crystal structures of CHIRAL and CLUMSY as challenges for synthesis and crystal-structure determination.

Lattice-vibrational entropies were also calculated and the corresponding contributions to thermal vibrational entropy of the different crystal structures were estimated. A larger entropy generally goes together with a larger cell volume and hence a smaller lattice energy for a given isomer. Entropy differences among computational polymorphs can sometimes compete with packing energy differences in the free energies at 300 K, indicating that lattice-vibrational entropy estimates are worth including in crystal structure prediction calculations. Although the entropy differences among competing virtual structures are small, so are the lattice energy differences, as also found experimentally for observed polymorphs.

There are obviously similarities between the crystal-packing problem and the protein-folding problem. Both involve delicate balances between attractions and

repulsions at the atomic level, between packing energy and entropic contributions to the free energy, and between thermodynamic and kinetic factors. If anything, the protein-folding problem may be even more difficult since it involves interactions between different side chains rather than between identical molecules. Moreover, to allow the flexibility necessary for function, the packing density in a protein molecule must be somewhat less than in a crystal, so even more structures must be attainable within a narrow energy range. On the other hand, vastly more effort is being invested in the protein folding problem because of its bio-medical importance. As time passes, it will be interesting to compare progress in these two areas.

REFERENCES

[1] A. I. Kitaigorodskii, 'Organic Chemical Crystallography', Consultants Bureau, New York, 1961 (English translation of the Russian original published by Press of the Academy of Sciences of the USSR, Moscow, 1955).
[2] A. I. Kitaigorodsky, 'Molecular Crystals and Molecules', Academic Press, New York, 1973.
[3] J. D. Dunitz, A. Gavezzotti, *Acc. Chem. Res.* **1999**, *32*, 677.
[4] D. Hilbert, 'Mathematische Probleme', Göttinger Nachrichten **1900**, 253. (English translation from M. Senechal, 'Quasicrystals and geometry', Cambridge University Press, 1995, p. 32).
[5] G. Polya, *Z. Kristallogr.* **1924**, *60*, 278.
[6] C. H. MacGillavry, 'Symmetry Aspects of M. C. Escher's Periodic Drawings', Oosthoek, Utrecht, 1965.
[7] A. Gavezzotti, *Atti Accad Lincei*, *Memorie Classe di Scienze Fisiche*, *Matematiche Naturali* **1976**, *XIII*, 107.
[8] J. D. Dunitz, G. Filippini, A. Gavezzotti, *Tetrahedron* **2000**, in press.
[9] A. Gavezzotti, G. Filippini, *J. Am. Chem. Soc.* **1995**, *117*, 12299.
[10] J. Bernstein, A. T. Hagler, *J. Am. Chem. Soc.* **1978**, *100*, 673.
[11] J. Bernstein, 'Conformational Polymorphism, in Organic Solid State Chemistry', Ed. G. R. Desiraju, Elsevier, Amsterdam, 1987, pp. 471–518.
[12] J. D. Dunitz, J. Bernstein, *Acc. Chem. Res.* **1995**, *28*, 193.
[13] W. C. McCrone, 'Polymorphism, in Physics and Chemistry of the Organic Solid State', Eds. D. Fox and M. M. Labes, A. Weissberger, Interscience, New York, 1965, Vol. 2, pp. 726–767.
[14] B. P. van Eijck, A. L. Spek, W. T. M. Mooij, J. Kroon, *Acta Crystallogr.*, *Sect. B* **1998**, *54*, 291.
[15] A. Gavezzotti, *Acta Crystallogr. Sect. B* **1996**, *52*, 201.
[16] A. Gavezzotti, *Crys. Rev.* **1998**, *7*, 5 (see in particular p. 18).
[17] A. Gavezzotti, *J. Am. Chem. Soc.* **1983**, *105*, 5220.
[18] A. Gavezzotti, *J. Am. Chem. Soc.* **1985**, *107*, 962.
[19] A. Gavezzotti, G. Filippini, *J. Phys. Chem.* **1994**, *98*, 4831.
[20] C. M. Gramaccioli, G. Filippini, *Chem. Phys. Lett.* **1984**, *108*, 585.
[21] J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer, D. E. Williams, *Acta Crystallogr.*, *Sect. B*, **2000**, *56*, 697.
[22] A. Gavezzotti, Zip-Promet, Program for the Generation of Crystal Structures from Molecular Structure, University of Milan, 1999.
[23] A. Gavezzotti, Minopec, A Program for the Optimization of Lattice Energies, University of Milan, 1999.
[24] B. P. van Eijck, W. T. M. Mooij, J. Kroon, *Acta Crystallogr., Sect. B* **1995**, *51*, 99; B. P. van Eijck, J. Kroon, *J. Comput. Chem.* **1999**, *20*, 799.
[25] C.-H. Lee, S. Liang, T. Haumann, R. Boese, A. de Meijere, *Angew. Chem.*, *Int. Ed.* **1993**, *32*, 559.
[26] A. Gavezzotti, *J. Chem. Soc.*, *Perkin Trans. 2*, **1995**, 1399.
[27] M. F. Richardson, Q.-C. Yang, E. Novotny-Bregger, J. D. Dunitz, *Acta Crystallogr.*, *Sect. B* **1990**, *46*, 653.
[28] J. D. Dunitz, G. Filippini, A. Gavezzotti, unpublished work.
[29] R. Boese, D. Blaser, W. E. Billups, M. M. Haley, A. H. Maulitz, D. L. Mohler, K. P. C. Vollhardt, *Angew. Chem. Int. Ed.* **1994**, *33*, 313.
[30] A. I. Kitaigorodskii, as quoted in [1], p. 414 (with minor geometrical corrections).
[31] D. Bright, I. E. Maxwell, J. de Boer, *J. Chem. Soc.*, *Perkin Trans. 2*, **1973**, 2101.

Received June 2, 2000