

1

Geometry and Energetics

1.1

Supramolecular Interactions: Energetic Considerations

Angelo Gavezzotti

“God always geometrizes”
(attributed to Plato, showing, if confirmed, that even great philosophers can be badly mistaken)

1.1.1

Introduction

In science as well as in everyday life, things are seldom as they appear to the unaided eye. Sight, that privileged sense so cherished by humans, leading as it does to the concepts of pleasure and beauty, is but of limited use in scientific proceedings, especially those connected with the behavior of matter at a molecular level. At a higher level of elaboration, geometry can be called for in the interpretation of objective phenomena, with some progress, especially in structural chemistry, where the geometrization of structure is of considerable help in the description of molecular-size objects. Chemistry, however, cannot be entirely built upon either sight or geometry. Electrons and nuclei act under the laws of quantum mechanics under well specified Hamiltonian operators, all around the key observable quantity, interelectronic and internuclear energy. Macroscopic systems span a space of dynamic variables, velocities and momenta, and evolve under the firm guidance of the two principles of thermodynamics, the most unnegotiable of all physical laws. There is very little solid chemistry without quantitative estimates of energy and entropy. This chapter is designed to outline the principles of the quantitative energetic evaluation of intermolecular interactions, the basis of nonreactive chemical phenomena occurring among organic molecules in condensed phases.

1.1.2

Enthalpy1.1.2.1 **The Quantistic Approach: Molecular Orbital (MO) Theory**

The fundamental equation of quantum mechanics is the time-dependent Schrödinger equation:

$$H \Psi = -\hbar/i(\partial \Psi / \partial t) \quad (1.1.1)$$

where H is the hamiltonian operator, \hbar is Planck's constant, and i is the imaginary unit ($i = \sqrt{-1}$). When the potential energy is independent of time, the wavefunction Ψ factorizes into a time-dependent part, $f(t)$, and a time-independent, position-dependent part, $\psi(\mathbf{x})$, so that:

$$H \psi(\mathbf{x}) = E \psi(\mathbf{x}) \quad (1.1.2)$$

$$\Psi(\mathbf{x}, t) = f(t) \psi(\mathbf{x}) \quad (1.1.3)$$

After normalization of the wavefunction, multiplying on the left by $\psi^*(\mathbf{x})$, and integrating over all space, one gets the expectation value of the electronic energy:

$$\langle E \rangle = \int \psi^*(\mathbf{x}) H \psi(\mathbf{x}) d\tau / \int \psi^*(\mathbf{x}) \psi(\mathbf{x}) d\tau = \int \psi^*(\mathbf{x}) H \psi(\mathbf{x}) d\tau \quad (1.1.4)$$

A hydrogen atom consists of a nucleus of unit positive charge and one electron. The quantum mechanical equation for a hydrogen atom is [1]:

$$-\hbar^2 / (2m_e) \nabla^2 \psi - (e^2/r) \psi = E \psi \quad (1.1.5)$$

where the first term corresponds to the kinetic energy part of the Hamiltonian, and the second represents the Coulombic potential between nucleus and electron. ∇^2 is the second derivative operator. This equation is the only one that can be exactly solved in quantum chemistry. For helium, one gets:

$$\begin{aligned} & -\hbar^2/2m_e \nabla_1^2 \psi - (2e^2/r_1)\psi - (\hbar^2/2m_e)\nabla_2^2 \psi - (2e^2/r_2)\psi + (e^2/r_{12})\psi \\ & = H_1\psi + H_2\psi + (e^2/r_{12})\psi = E \psi \end{aligned} \quad (1.1.6)$$

where now the equation contains one cross term for the Coulombic interaction between the two electrons. Because of this last term, the analytical solution of the above equation is impossible.

For a polyatomic molecule, nuclei are considered motionless, because the motion of electrons is on a much faster timescale than the motion of nuclei (the Born–Oppenheimer approximation). The corresponding Hamiltonian is:

$$H = \sum_k H(k) + e^2 \sum_{k>j} 1/r_{kj} \quad (1.1.7)$$

where $H(k)$ is the part for electron k , kinetic energy plus coulombic interactions with all the nuclei, and the second summation corresponds to the electron–electron Coulombic potential. Since an analytical solution of the Schrödinger equation is impossible even for a helium atom, the case with a hamiltonian like (1.1.7) is plainly hopeless.

Solution of the Schrödinger equation provides two main pieces of information: the total electronic energy of the system (the repulsion energy between motionless nuclei can be simply calculated by a Coulomb sum), and the wavefunction $\Psi(x, y, z, t)$, which in turn gives the electron density $\rho(x, y, z)$ and a formal answer to the question ‘where are the electrons?’, because the quantity:

$$P(d\tau) = [\Psi\Psi^*] dx dy dz \quad (1.1.8)$$

is interpreted as the probability, P , that an electron be found in the infinitesimal volume $dV = dx dy dz$. Ψ^* is the complex conjugate of Ψ .

In order to solve Eq. (1.1.2) with Hamiltonian (1.1.7), a large number of approximations are needed. Enter the favorite son of quantum chemistry, the molecular orbital (MO) method [1]. As appears from Eq. (1.1.4), a great many integrals are needed, so the wavefunction must be written in an easily integrable form. The nodal properties, crucial for the energetic considerations, just as the number of nodes relates to the energy of a vibrating string, must be taken into account: the spherical harmonics, $Y_{lm}(\theta, \phi)$, are used to take care of this. In MO theory, the radial part of atomic orbitals, $R_i(r)$, is written as a combination of gaussian basis functions $g(r)$, complete atomic orbitals χ_i are written as products of the radial and angular parts, and molecular orbitals ϕ_i are then obtained as linear combinations of atomic orbitals (LCAO):

$$g(r) = N \alpha^{(2n+1)/4} r^{n-1} \exp(-\alpha r^2) \quad (1.1.9)$$

$$R_i(r) = \sum_j a_{ij} g_j \quad (1.1.10)$$

$$\chi_i = N' R_i(r) Y_{lm}(\theta, \phi) \quad (1.1.11)$$

$$\phi_i = \sum_j c_{ij} \chi_j \quad (1.1.12)$$

in which r is the distance of the electron from the nucleus, α is called the exponent of the gaussian, n is 1, 2 or 3 for s, p or d functions respectively, and N and N' are normalization constants.

The total wavefunction Φ for a molecular system is then taken as a product of molecular orbitals, each of which contains the coordinates of one electron only, i.e., Φ is a product of one-electron orbitals. This is a very severe approximation, because this wavefunction cannot describe the effects of electron correlation, or the simultaneous displacement, or other simultaneous change in properties, of many electrons at a time.

Spin is a further complication. According to the Pauli principle, acceptable wavefunctions must change sign on formal exchange of the coordinates of two electrons [1]. This is taken care of by writing the total wavefunction in the form of a determinant (the Slater determinant), in which each MO is multiplied by an appropriate spin wavefunction.

The solution of Eq. (1.1.2) is done in an iterative, self-consistent way using the variational principle: whenever the true wavefunction is approximated by some incomplete function that depends on a number of parameters, the expectation value of the energy, Eq. (1.1.4), is higher than the expectation value that competes to the exact wavefunction. The procedure requires a lot of mathematical manipulation, but the problem has been solved once and for all and the task is nowadays performed by very efficient black-box computer packages.

An MO calculation provides, among many others, the following chemically useful items: (i) the energies of the molecular orbitals; these do not have a real physical meaning, but are sometimes useful, being an estimate of the energy levels of separate electron pairs in the molecule; (ii) the expectation value, or the total electronic energy of the molecule; (iii) the electron density, $\rho(x, y, z)$. For the purposes of supramolecular chemistry, total energies can be used to evaluate the intermolecular interaction energies in pairs or clusters of molecules, while the electron density determines the molecular electrostatic potential and hence the cohesive properties of the molecule and its molecular recognition abilities:

$$\rho(x, y, z) = \sum_r \sum_s P_{rs} \chi_r(x, y, z) \chi_s(x, y, z) \quad (1.1.13)$$

where P_{rs} is a sum of products of LCAO coefficients in Eq. (1.1.12). Energies obtained by the methods so far described are called Hartree–Fock (HF) energies. They represent the kinetic energies of electrons and the electrostatic effects between electrons and nuclei, but cannot take into account electron correlation, because of the use of one-electron orbitals. This limitation forbids the treatment of dispersion energy, a crucial part of the intermolecular potential. To enhance the scope of MO methods beyond HF, configuration interaction can be used. In the basic HF treatment, the total wavefunction is represented by one Slater determinant, but one could simultaneously take into account all the possible electronic configurations, the ground state with no excitation, plus single and double or even triple excitations, taking the total wavefunction as a linear combination of all the resulting Slater determinants. The final equations become very complicated, but are still manageable.

The inclusion of excited states is one way of representing electron correlation. In the Moeller–Plesset (MP) approximation, at the n th level (MP n) all the n -fold excited configurations are taken into account, and the total energy results as:

$$E = E^0 + E^1 + E^2 + E^3 + E^4 + \dots \quad (1.1.14)$$

The MP0 term is the sum of the energies of the molecular orbitals, the MP1 energy is the HF energy, and the energies that include upper terms in the expansion are called MP2, MP3, MP4... energies and represent electron correlation.

1.1.2.2 The Quantistic Approach: Density Functional Theory (DFT)

The foundation of DFT [2–4] is the Hohenberg–Kohn theorem: the energy of a system of electrons in an external potential $v(R)$ is a functional of the electron density (a functional is a mathematical entity that has a given value for each given value of another function), and the ground state energy of a many-electron system is uniquely determined by the ground-state density. In the Kohn–Sham assumption, a local potential, $V_{KS}(R)$, can be introduced for a molecular electron cloud in the field of the nuclei, such that a system of noninteracting electrons moving in the $V_{KS}(R)$ field will have the same density as the exact density of the interacting electron system. V_{KS} has terms for the electron–nuclei and electron–electron coulombic interaction, and a further term which incorporates the exchange [3] and electron correlation effects:

$$V_{KS}(R) = V_{eZ}(R) + V_{ee}(R) + V_{exch}(R) \quad (1.1.15)$$

At this point, the Kohn–Sham Hamiltonian operator can be written and the expectation value determined (compare with the above proceedings for the MO theory):

$$H_{KS} = -1/2\nabla^2 + V_{KS}(R) \quad (1.1.16)$$

$$\int \Phi^* H \Phi \, d\tau = \langle E_{KS} \rangle \quad (1.1.17)$$

$$E_{KS} = T_{KS} + U_{eZ} + U_{ee} + E_{exch,KS} \quad (1.1.18)$$

where T_{KS} is the kinetic energy, U_{eZ} is the electron–nuclei attraction, U_{ee} is the electron–electron repulsion, and $E_{exch,KS}$ is the exchange energy over the Kohn–Sham orbitals. The crucial point is the following: the exchange energy $E_{exch,KS}$ is replaced by an effective contribution $E_{ex,eff}$ that also includes correlation effects as well as possible, and whose functional form $\epsilon_{ex}(R)$ can be more or less empirically adapted to the purpose. This effective correction energy has a term in the difference between the true kinetic energy and the kinetic energy of the noninteracting electrons, and a term in the difference between the total electron interaction energy and the coulombic electron–electron repulsion. The latter term incorporates the correlation energy, as much as the functional $\epsilon_{ex}(R)$ is indeed a realistic energy density. In the local density approximation, LDA, one takes the exchange–correlation energy of an electron gas of uniform density, equal to the local density in the molecular system at point R : a drastic approximation. Other more sophisticated approaches have been proposed, hence the various names (B, BP, B-LYP, etc.).

1.1.2.3 The Quantistic Approach: the Crystal Orbital Method

So far, only energies of isolated molecules have been considered. In a crystal, [5] the translational invariance of the potential energy V at a generic point r in real space is

reflected in a translational invariance of the quantum problem, which is, as usual, the solution of a steady state Schrödinger equation:

$$V(\mathbf{r}) = V(\mathbf{r} - \mathbf{R}) \quad (1.1.19)$$

$$H \psi(\mathbf{r}) = E \psi(\mathbf{r}) \text{ or } H \psi(\mathbf{r} - \mathbf{R}) = E \psi(\mathbf{r} - \mathbf{R}) \quad (1.1.20)$$

where \mathbf{R} is any translational vector. The eigenfunctions which are proper solutions of Eq. (1.1.20) must be symmetry-adapted to translation; they are called Bloch functions:

$$\Phi(\mathbf{r} + \mathbf{R}, \mathbf{k}) = \exp(i\mathbf{k}\mathbf{R}) \Phi(\mathbf{r}, \mathbf{k}) \quad (1.1.21)$$

\mathbf{k} is a vector whose components have the dimensions of a reciprocal length, and the dot product in the exponential is a pure number. Since $\exp(i\mathbf{k}\mathbf{R})$ is a wave, \mathbf{k} expresses its wavelength (hence the name, wave vector).

A specification of the actual form of Bloch functions in terms of atomic orbitals is needed. They can be written as a product of a plane wave and a function, $\Omega(\mathbf{r}, \mathbf{k})$, with the translational periodicity of the lattice. With the Schrödinger problem written in \mathbf{k} -dependent form, the solutions take the form of a linear combination of Bloch functions, whose coefficients are to be determined. For all the atomic orbitals $\chi_j(\mathbf{r} - \mathbf{r}_j)$ in the unit cell, each centered at position \mathbf{r}_j , in a crystal with N unit cells, a combination which has the same translational periodicity as the lattice is the replacement of a single AO in the reference cell with a sum over all translationally equivalent AOs in the extended crystal. The final form of the crystal orbitals is then:

$$\Phi(\mathbf{r}, \mathbf{k}) = N^{-1/2} \sum_{\mathbf{R}} \exp[i\mathbf{k}\mathbf{R}] \chi_{j\mathbf{R}} \quad (1.1.22)$$

The solution of Eq. (1.1.20) with the hamiltonian of Eq. (1.1.7) and under the Bloch conditions (1.1.22) is conceptually similar to the solution under the LCAO-MO assumption (1.1.12). A major difference is that while for the molecular problem one solution is sufficient, the crystal problem is a function of \mathbf{k} and therefore the variational problem must be solved a number of times equal to the number of sampling points in the independent part of \mathbf{k} -space (the first Brillouin zone). These computational difficulties limit the applicability of the crystal orbital method to rather small molecules and unit cells, the urea crystal being probably the limit. Moreover, the crystal orbital method cannot be applied with a proper consideration of correlation energy because of computational limitations.

For a crystal, orbital energies are a function of \mathbf{k} and hence $E(\mathbf{k})$ plots are obtained, called energy bands. Like molecular orbitals, bands are partly filled with electrons and partly empty. The energy difference between the Fermi level (highest occupied point of a band) and the energy of the lowest unoccupied band is called the band-gap, the equivalent of the HOMO–LUMO gap in an isolated molecule. Metals have half-filled electronic energy bands and a zero band-gap, so

they are good conductors, while organic crystals have wide band-gaps and are therefore semiconductors or electric insulators.

The physical meaning of k -dependence can be further explained by comparing with the molecular case. In a molecule, atomic orbitals combine into molecular orbitals, whose energy increases with the number of nodes. In an infinite periodic system, like a crystal, one gets, in principle, an infinite number of crystal orbitals, which merge into a continuum represented in the continuous nature of the energy bands (which however, in practice, are sampled at a finite number of points in the Brillouin zone). Thus, k labels the different crystal orbitals and counts their number of nodes. Since k is a vector in crystal reciprocal space (the same reciprocal space used by X-ray crystallographers), how bands evolve in k depends on the nature of the electron density along the corresponding directions in crystal space.

1.1.2.4 The Classical Approach: Vibrational and Nonbonded ("Force Field") Energies

The Born–Oppenheimer approximation uncouples electron and nuclear motion. The latter concerns massive (at least, relative to electrons) bodies, and much lower velocities: while the formal velocity of an electron may approach the speed of light, a molecule in the gas phase travels at about the speed of a supersonic jet plane. While electronic energies must be calculated by quantum mechanics, nuclear motions are more easily described in a classical framework.

For a molecule with N nuclei, there are $3N$ degrees of freedom, with three ascribed to rigid-body translation, three to rotation, and $3N-6$ to vibration, or the change of the reciprocal positions of the nuclei. The classical potential energy of such vibrations can be written as:

$$E(\text{pot,vib}) = \sum_i 1/2 k_i (g_i - g_i^\circ)^2 + \sum_{i,j} 1/2 k_{ij} (g_i - g_i^\circ) (g_j - g_j^\circ) + \sum_n f(\tau_n) \quad (1.1.23)$$

where the k s are the force constants, provided by molecular spectroscopy or by empirical fitting; the g s are the actual values of some internal bond stretching or bond bending coordinates, and g° is the corresponding reference, or 'strainless' value, obtained usually by statistical studies of spectroscopic or diffraction experiments. $f(\tau)$ is some function of torsional angles. In addition, one must have some way of representing the soft attractive potential that arises from interaction between atoms not bound to one another and not bound to the same atom, as well as the sharp repulsion that arises when these atoms come into too close contact. This nonbonded potential energy is usually given the empirical form: [6–9]

$$E(\text{nonbonded}) = \sum_{i,j} f(R_{ij}) + E(\text{Coulombic}) \quad (1.1.24)$$

$$f(R_{ij}) = -AR_{ij}^{-n} + BR_{ij}^{-m} + C \exp(-DR_{ij}) \quad (1.1.25)$$

where R_{ij} is the distance between two atomic nuclei, A , B , C , D ..., n , m ... are empirical disposable parameters. Typical values are $n=6$ and $m=12$. The Coulombic

term is supposed to represent the direct coulombic interactions between polarized regions of the electron distribution of the molecule; in the simplest case, it takes the point-charge form: [10]

$$E(\text{Coulombic}) = 1/(4\pi\epsilon^0) q_i q_j R_{ij}^{-1} \quad (1.1.26)$$

where the q_i s are point-charge parameters. In an immediate extension, these same functional forms can be used for intermolecular interactions, the idea being that the interaction energy between the two phenyl rings in 1,3-diphenylpropane has the same origin as that between two benzene rings in the 1,3-diphenylpropane crystal. Table 1.1.1 shows some typical stretching force constants, Table 1.1.2 some typical nonbonded functions. It may at once be seen that a 0.1 Å displacement in the region around the minimum causes an energy rise of 5–10 kJ mol⁻¹ for the former, and of 0.5–1 kJ mol⁻¹ for the latter. There is an order-of-magnitude difference between intra- and intermolecular stretchings.

Table 1.1.1 Some typical force constants over internal molecular coordinates (units of N m⁻¹).

C–C stretching, alkanes	430–450
Csp ³ –Csp ²	500
C–C stretching, benzene	625
C–C double bond stretching, olefins	910
C–H stretching, alkanes	450–470
C=O stretching	800

Table 1.1.2 UNI force field [8, 9] potential parameters: A, B, C for $A \exp(-BR) - CR^{-6}$ Distances in Å, energies in kJ mol⁻¹.

		A	B	C	R°	Well depth
H	H	24158.4	4.010	109.20	3.36	-0.042
H	C	120792.1	4.100	472.79	3.29	-0.205
H	N	228279.0	4.520	502.08	2.98	-0.394
H	O	295432.3	4.820	439.32	2.80	-0.505
H	F	64257.8	4.110	248.36	3.29	-0.110
H	S	268571.0	4.030	1167.34	3.35	-0.458
H	CL	292963.7	4.090	1167.34	3.30	-0.501
C	C	226145.2	3.470	2418.35	3.89	-0.387
C	N	491494.5	3.860	2790.73	3.49	-0.851
C	O	393086.8	3.740	2681.94	3.61	-0.674
C	F	196600.9	3.840	1168.75	3.50	-0.350
C	S	529108.6	3.410	6292.74	3.96	-0.909
C	CL	390660.1	3.520	3861.83	3.83	-0.678

Table 1.1.2 (continued)

		A	B	C	R°	Well depth
N	N	365263.2	3.650	2891.14	3.70	-0.629
N	O	268571.0	3.860	1522.98	3.49	-0.464
N	F	249858.9	3.930	1277.90	3.43	-0.435
N	S	630306.9	3.590	5576.76	3.75	-1.108
O	O	195309.1	3.740	1334.70	3.61	-0.336
O	F	182706.1	3.980	868.27	3.39	-0.320
O	S	460909.4	3.630	3790.70	3.71	-0.801
O	CL	338297.3	3.630	2782.36	3.71	-0.588
F	F	170916.4	4.220	564.84	3.20	-0.293

The conformation of a gas-phase molecule can be derived by minimizing the strain energy Eqs. (1.1.23)–(1.1.26). A powerful simulation tool for the collective behavior of a molecular ensemble is molecular dynamics (MD) [11]. In the classical approach, the total instantaneous energy of a molecular system, $E(\text{tot})$ is again a sum of terms as in Eqs. (1.1.23)–(1.1.26). Note that the above functional forms assume that the electronic structure is not significantly perturbed, that is, $E(\text{electronic}) = \text{constant}$.

Forces $F_{i,k}$ (k th cartesian component of the position vector of atom i) can be calculated easily, given the simple functional forms of vibrational and nonbonded potentials, as their first derivatives. The time evolution of such a system can be described using the equations of motion of classical Newtonian mechanics, $F_{i,k} = m_i(d^2x_{i,k}/dt^2)$. Given a reasonable starting configuration $(x_{i,k})^\circ$, integration of the above differential equation gives the trajectories of all atoms in the molecule. Of course the integration cannot be carried out analytically, but is carried out by fast and efficient numerical iterative techniques.

The total kinetic energy of the multimolecular system is the sum of the kinetic energies of all atoms, and can be equated to the equipartition value, in the key link between molecular motion and temperature:

$$E(\text{kin}) = \sum_i 1/2 m_i v_i^2 = 1/2 k_B T (n_{\text{dof}}) = 1/2 k_B T (3n_{\text{atoms}}) \quad (1.1.27)$$

where m_i and v_i are the mass and velocity of the i th atom.

Consider now the statistical mechanics definition of internal pressure:

$$P = 2/(3V) [E(\text{kin}) - \Xi] \quad (1.1.28)$$

$$\Xi = -1/2 \sum_k \sum_l R_{k,l} F_{k,l} \quad (1.1.29)$$

V is the volume of the system, and the quantity Ξ is called the virial, a sum of products of the distance between any two atoms k and l in the system and of the force acting between them. Distances are always positive numbers so the virial is positive

when forces are negative (that is attractive). For an ideal gas, $E(\text{kin}) = 3/2RT$ per mole and since forces are all zero $\Xi = 0$ and $P = RT/V$; for real systems, including condensed phases, the attractive forces between molecules dictate the virial to subtract momentum from molecular action towards the surroundings. Equations (1.1.23)–(1.1.29) define the total energy of a molecular ensemble as a function of time, temperature and pressure: molecular dynamics samples the phase space and allows a calculation of complete $P/V/T$ phase diagrams.

1.1.2.5 Semi-classical Approaches: the SCDS-Pixel Method

The quantum mechanical approach cannot be used for the calculation of complete lattice energies of organic crystals, because of intrinsic limitations in the treatment of correlation energies. The classical approach is widely applicable, but is entirely parametric and does not adequately represent the implied physics. An intermediate approach, which allows a breakdown of the total intermolecular cohesion energy into recognizable coulombic, polarization, dispersion and repulsion contributions, and is based on numerical integrations over molecular electron densities, is called semi-classical density sums (SCDS) or more briefly 'Pixel' method. [12–14]

In the Pixel formulation, the basic concept is the electron density unit, or e-pixel. Let q_k be the electron density, calculated using a MO wavefunction for molecule A, in a volume V_k centered at point $(k) = [x_k, y_k, z_k]$. Each of the e-pixels is assigned a charge $q_k = \rho_k V_k$. Super-pixels are then formed by condensing $n \times n \times n$ original pixels, n being called the condensation level. In addition, molecule A has nuclei of charge Z_j at points $(j) = [x_j, y_j, z_j]$. A second molecule B has its e-pixels with $q_i = \rho_i V_i$, at positions $(i) = [x_i, y_i, z_i]$, and nuclei of charge Z_m at points $(m) = [x_m, y_m, z_m]$. Each molecule in the computational box is thus represented by a set of e-pixel charges, typical numbers being 10 000 for benzene to 25 000 for anthracene, plus a set of nuclear charges.

For the calculation of Coulombic energies, the electrostatic potential Φ_i generated by molecule A at point (i) of the charge density of molecule B, and the potential Φ_m generated by molecule A at nucleus m of molecule B, are calculated, and the total electrostatic potential energy between the two molecules is the sum of the electrostatic energies at points (i) , $E_i = q_i \Phi_i$, and at points (m) , $E_m = Z_m \Phi_m$.

Calling ϵ_i the total electric field exerted by surrounding molecules at pixel i , α_i the polarizability at pixel i , and μ_i the dipole induced at pixel i by that field, the linear polarization energy is:

$$E_{\text{POL},i} = -1/2 \mu_i \epsilon_i = -1/2 \alpha_i \epsilon_i^2 \quad (1.1.30)$$

In order to assign pixel polarizabilities, each pixel has to be first assigned to a particular atom in the molecule. Let p be the number of atoms for which the nucleus–pixel distance is smaller than the atomic radius. If $p=1$, the charge pixel is within one atomic sphere only, and it is assigned to that atom. If $p > 1$, the pixel is assigned to the atom from which the distance is the smallest fraction of the atomic radius. If $p = 0$, the pixel is assigned to the atom whose atomic surface is nearest.

Atomic polarizabilities are taken from standard repertories, and pixel polarizabilities are calculated from them as $\alpha_i = (q_i/Z_{\text{atom}}) \alpha_{\text{atom}}$.

Polarization contributions at very short distance between polarizer and polarized pose divergence problems, because of the R^{-3} dependence of the polarization energy, and must be somehow damped. Electric fields in molecular crystals are mostly of the order of 10^{10} V m^{-1} , plus a small number in the 10^{10} – 10^{13} V m^{-1} range, plus a very small number of even higher ones; these high-field contributions are physically unrealistic, resulting from fortuitous short distances in the overlapping density meshes. The damped polarization energy at pixel i is

$$E_{\text{POL},i} = -1/2 \alpha_i [\epsilon_i d_i]^2 \text{ for } \epsilon < \epsilon_{\text{max}}, d_i = \exp(-\epsilon_i / (\epsilon_{\text{max}} - \epsilon_i)) \quad (1.1.31)$$

and $E_{\text{POL},i} = 0$ for $\epsilon > \epsilon_{\text{max}}$. ϵ_{max} is an adjustable empirical parameter in the formulation, set at $150 \times 10^{10} \text{ V m}^{-1}$. The total polarization energy at a molecule is the sum of polarization energies at each of its electron density pixels, $E_{\text{POL,TOT}} = \Sigma E_{\text{POL},i}$, while the total polarization energy in an ensemble of molecules is the sum of all A...B and B...A polarization energies. Polarization energies are not pairwise additive over atomic contributions.

Intermolecular dispersion energies are calculated as a sum of pixel–pixel terms in a London-type expression, involving the above defined distributed polarizabilities and an overall ‘oscillator strength’, E_{OS} . To avoid singularities (as before described) due to very short pixel–pixel distances in an inverse sixth-power formula, each term in the sum is damped, as it is shown here for the molecule A...molecule B interaction:

$$E_{\text{DISP,AB}} = E_{\text{OS}} (-3/4) \Sigma \Sigma f(R) \alpha_i \alpha_j / [(4\pi\epsilon^0)^2 (R_{ij})^6] \quad (1.1.32)$$

where $f(R)$ is the damping function, $f(R) = \exp[-(D/R_{ij} - 1)^2]$ for $R_{ij} < D$ and $f(R) = 1$ for $R_{ij} > D$. D is an adjustable empirical parameter, set at 3.50 \AA . E_{OS} was originally approximated by the molecular ionization potential, for very small molecules. In complex polyatomic molecules, E_{OS} can be taken as the energy of the highest occupied molecular orbital, the interacting electrons being usually the peripheral ones and hence roughly at the HOMO energy level. This assumption works well with small molecules containing C, H, N or O atoms, but fails for example for heavily fluorinated aromatic compounds, when the interacting electrons belong with the fluorine atoms, whose ionization potential is much lower than the energy of the HOMO, which is usually a π -type molecular orbital. The London oscillator strength E_{OS} for the interaction between any two molecules can then be calculated taking into account the different nature of the interacting electrons, as a weighted sum of atomic ionization potentials, weights being the atom–atom overlap integrals. This procedure amounts in fact to taking different dispersion energy coefficients according to the different kinds of approaching atomic basins.

The repulsion energy, E_{REP} , is modeled as proportional to intermolecular overlap. The total overlap integral between the charge densities of any two molecules, S_{AB} , is calculated by numerical integration over the original uncondensed densities, and is

then subdivided into contributions from pairs of atomic species, S_{mn} , using the assignment of pixels to atomic basins. For each $m-n$ pair the repulsion energy is evaluated as

$$E_{\text{REP},mn} = (K_1 - K_2 \Delta\chi_{mn}) S_{mn} \quad (1.1.33)$$

where $\Delta\chi_{mn}$ is the corresponding difference in Pauling electronegativity. The values of K_1 and K_2 were optimized at 4800 and 1200 respectively (for energies in kJ mol^{-1} with electron densities in $\text{electrons \AA}^{-3}$). The rationale behind this approach is that when atoms of different electronegativity meet, a larger reorganization of the electron density occurs, and the overlap repulsion must be smaller. The total repulsion energy is then the sum over all $m-n$ pairs.

The total intermolecular interaction energy is then:

$$E_{\text{TOT}} = E_{\text{COUL}} + E_{\text{POL}} + E_{\text{DISP}} + E_{\text{REP}} \quad (1.1.34)$$

For a crystal, a number of symmetry related molecules (nuclear positions plus charge density pixels) are generated around a reference molecule, and all the terms of Eq. (1.1.34) for all the molecular pairs are evaluated. Their sum is the lattice energy of the crystal.

1.1.2.6 Supramolecular Energies

The internal energy per mole of a chemical system is the sum of all energies, electronic, vibrational, and kinetic, possessed by one mole of molecules at a given temperature, pressure and volume. Its thermodynamic symbol is U , and the relationship to enthalpy, H , is given by $H = U + PV$. Chemists use enthalpies because for any process, dH is equal to the heat exchanged, dq , without contributions from volume work.

Consider first the simple molecular recognition process in which molecule A meets molecule B and a molecular complex is formed. In the quantistic approach the energy variation upon complex formation can be written as:

$$\begin{aligned} \Delta E(\text{compl}) &= \Delta E(\text{electronic}) + \Delta E(\text{pot,vibr}) + \Delta E(\text{kin}) \approx \\ &\Delta E(\text{electronic}) + \Delta E(\text{zero point, vibr}) \end{aligned} \quad (1.1.35)$$

$\Delta E(\text{electronic})$ must be calculated by quantum chemical methods by a difference between the total energy of the complex and the sum of the energies of the separated molecules. At the formally zero temperature of the quantum calculation, kinetic energies are zero, and vibrational terms are the zero-point contributions (i.e. those coming from vibrational levels of quantum number zero). These differences are very small and are often neglected, although they can be of the same order of magnitude as the difference in electronic energies.

Consider now an ensemble of molecules in the classical approach (force field + MD) to molecular energies, and the sum:

$$E(\text{pot,vibr}) + E(\text{nonbonded}) + E(\text{kin}) = U - U^\circ \quad (1.1.36)$$

where the terms are taken as averages of the sampling of phase space during an MD simulation, and U° is an unknown zero reference term, anyway unimportant since only energy differences are meaningful. Sum (1.1.36) is the total internal energy of the molecular system, and is conceptually equivalent to the statistical mechanics internal energy, or to the classical thermodynamics internal energy:

$$U - U^\circ = -(\partial \ln Q / \partial \beta) \quad \text{statistical mechanics, ideal gas} \quad (1.1.37)$$

$$\Delta U = q + w \quad \text{classical thermodynamics} \quad (1.1.38)$$

where Q is a molecular partition function, and q and w are heat and work exchanges in a thermodynamic process. In molecular dynamics, U is an average over the time evolution of the value of energy function, Eq. (1.1.36); the classical approach has the advantage that the simulation also shows the actual dynamics of the process, and can be made temperature- and pressure-dependent.

Other interesting properties, related to molecular cohesion in condensed phases, are the enthalpies of vaporization and sublimation. In the classical MD approach, the vaporization enthalpy of a liquid or of a solid is:

$$\Delta H(\text{vap}) = \Delta U(\text{vap}) + P\Delta V(\text{vap}) = \Delta U(\text{vap}) + PV(\text{g}) = \Delta E(\text{intra}) + \Delta E(\text{inter}) + RT \quad (1.1.39)$$

where the volume of the condensed phase is negligible with respect to the volume of the gas, considered ideal. Kinetic energies do not appear because $E(\text{kin})$ is the same for any phase at the same temperature. For rigid molecules, $\Delta E(\text{intra})$, the intramolecular potential energy difference between gas and condensed phase, is ≈ 0 ; the intermolecular potential in the gas is assumed equal to zero, so that the intermolecular potential energy difference, $\Delta E(\text{inter})$, is simply equal to the negative of the intermolecular energy in the condensed phase, so that for a crystal $\Delta H(\text{vap})$ is equal to $-E(\text{lattice}) + RT$. When the calculation is static, i.e. the lattice energy is derived by a single calculation on the equilibrium crystal structure, a somewhat different line of reasoning leads to $\Delta H(\text{subl}) = -E(\text{lattice}) - 2RT$. In practice these corrections are often absorbed into the parametrization and, conveniently, $\Delta H(\text{subl}) = -E(\text{lattice})$. Typical sublimation enthalpies of organic compounds are in the range 50–200 kJ mol⁻¹.

1.1.3

Entropy

1.1.3.1 Statistical and Classical Entropy

Contrary to what happens for mechanical systems, chemical systems do not reach equilibrium by minimizing their internal energy. The essential difference is that in

chemical systems molecular motion incessantly redistributes the available energies among the constituting molecules in a completely random fashion, and part of the driving force towards equilibrium comes from reaching the statistically most probable distribution of these energies. Entropy cannot be defined in terms of order and disorder, subjective concepts which cannot be defined in a proper way; entropy has to do with the distribution of available energies. This is evident from the statistical thermodynamic definition of entropy. Let M be the total number of molecules in a system, and m_i^* the number of molecules in a given quantum level ε_i at equilibrium. The statistical weight of the equilibrium distribution, W^* , and the statistical entropy, S , are given by

$$W^* = M! / (\prod m_i^*) \quad (1.1.40)$$

$$S = [k_B/N] \ln(W^*) \quad (1.1.41)$$

N here is the number of particles in the system, or Avogadro's number for one mole. W^* is also the number of different, indistinguishable ways in which the equilibrium distribution of energies can be obtained, or, in a quite similar way, in which the equilibrium distribution of spatial orientations can be obtained. Thus in a perfectly symmetrical crystal at zero temperature all molecules are in the ground energy level, and all molecules are oriented according to symmetry, so there is only one way of distributing energies and there is only one way of distributing molecular orientations; accordingly, the logarithm dependence makes entropy zero, independently of what the symmetry is (a P1 crystal with just translational symmetry has the same zero entropy as an Fdd2 crystal with 16 equivalent positions in the unit cell, as long as symmetry is obeyed in both). This is sometimes called the third principle of thermodynamics.

Further manipulations give the Boltzmann equilibrium distribution and the total entropy of an ideal gas:

$$m_i^* = M \exp(-\beta E_i) / Q; P_i^* = m_i^* / M \quad (1.1.42)$$

$$S = -k_B \sum [(m_i^* / M) \ln (m_i^* / M)] = -k_B \sum P_i^* \ln P_i^* \quad (1.1.43)$$

$$S = (U - U^0) / T + k_B \ln(Q) = -1/T [(1/Q)(\partial Q / \partial \beta)_{V}] + k_B \ln(Q) \quad (1.1.44)$$

The partition function Q has a translational, rotational and vibrational part, and so does entropy. The first two contributions can be calculated in a straightforward way, but the vibrational contribution to entropy requires a knowledge of the normal vibrational frequencies of the molecule or of the crystal.

The classical definition of entropy goes through the Clausius inequality:

$$dS > dq/T \quad \text{irreversible processes} \quad (1.1.45)$$

$$dS = dq/T \quad \text{reversible processes or equilibrium}$$

which, after rearranging, yield $(\partial S/\partial P)_T = -(\partial V/\partial T)_P$ and $(dS)_P = (dH/T)_P = (C_p/T)dT$, with access to entropy variations through measurable quantities like the expansion coefficient and the heat capacity.

1.1.3.2 Lattice Dynamics [15] and Lattice Vibration Frequencies

In the harmonic approximation for a molecular crystal, all atoms are oscillating about their equilibrium positions under the restraining action of a vibrational potential, V , which can be conveniently taken as just the intermolecular non-bonded potential of Eq. (1.1.25), and which obeys the translational invariance condition of Eq. (1.1.19). Let x_i be the displacement of atom i from its equilibrium position in the crystal, and consider all the atoms in the reference cell and all the atoms in the surrounding cells denoted by a real space vector \mathbf{R} . The mass-weighted force constants can be written as:

$$D_{ij,\mathbf{R}} = (m_i m_j)^{-1/2} [\partial^2 V / \partial x_{\text{ref},i} \partial x_{\mathbf{R},j}]^\circ \quad (1.1.46)$$

where the m_i s are atomic masses, and the zero superscript recalls that the derivatives are calculated at the equilibrium position. These force constants have the same lattice periodicity as the potential, so they can be treated just as the electronic orbitals were treated, i.e. they are periodicized through a Bloch expansion:

$$D_{ij}(\mathbf{k}) = \sum_{\mathbf{R}} \exp[i\mathbf{k}\mathbf{R}] D_{ij,\mathbf{R}} \quad (1.1.47)$$

The dynamic problem of vibrational spectroscopy must be solved to find the normal coordinates as linear combinations of the basis Bloch functions, together with the amplitudes and frequencies of these normal vibrations. These depend on \mathbf{k} , and therefore the problem must be solved for a number of \mathbf{k} -points to ensure an adequate sampling of the Brillouin zone. Vibrational frequencies spread in \mathbf{k} -space, just as the Bloch treatment of electronic energy gave a dispersion of electronic energies in \mathbf{k} -space. The number of vibrational levels whose energy lies between E and $E+dE$ is called the vibrational density of states. Vibrational contributions to the heat capacity and to the crystal entropy can be calculated by appropriate integrations over the vibrational density of states, just like molecular heat capacities and entropies are obtained by summation over molecular vibration frequencies.

Just as for crystal orbitals, \mathbf{k} is a reciprocal space vector that identifies the vibrational normal modes and counts the number of nodes in the corresponding combination of atomic vibrations. How the energy of a given vibration evolves in \mathbf{k} -space depends on the intermolecular potential restraints along that particular direction. In organic crystals, the forces are anyway moderate and no major anisotropies may arise – compare, for instance, with the case of an ionic crystal in which one finds alternations of positive and negative charges. In organic crystals lattice vibrational frequencies fall in a rather restricted range, something like 30–150 cm^{-1} , and even \mathbf{k} -spreading cannot bring about major differences. The vibrational density of states of organic crystals are all similar, and therefore the

corresponding vibrational entropies cannot be too different. Enthalpy differences between polymorphs must be small because the potentials are anyway weak and almost isotropic, and entropy differences follow suit, because the force constants (the second derivatives of these potentials) also cannot be too different. Typical calculations show lattice-dynamical entropy differences between computational polymorphs in a range of 10–15 J K⁻¹ mol⁻¹, implying a TΔS range at room temperature of less than 5 kJ mol⁻¹, just the same range as that of the enthalpy differences.

1.1.3.3 Entropy and Dynamic Simulation

The internal energy (or enthalpy) of a liquid or crystalline molecular ensemble can be easily evaluated by a MD simulation, because even a moderate sampling of phase space is sufficient to provide an adequate representation and a meaningful average. Usually 100 ns are sufficient for liquids at not too high temperatures, and even 10 ns may be enough for a crystal, whose chances to roam over phase space are much more restricted. In fact, for a crystal even a single calculation over a static model gives a rather accurate evaluation of the lattice energy. Entropy, on the other hand, is connected with the size of phase space actually available to the system and cannot be obtained as an average of some ensemble property over a limited simulation; in principle, all phase space should be swept and checked for availability, an obviously impossible task. The estimation of entropies by molecular dynamics simulations uses statistical approaches to the evaluation of work and heat exchanges, but the involved calculations are neither simple nor affordable in a routine way to the casual investigator. [16, 17]

1.1.4

Free Energy

1.1.4.1 Complexation and Evaporation/Sublimation

Gibbs free energy, $G = H - TS$, or Helmholtz free energy, $F = U - TS$, are the key quantities in the description of chemical equilibrium. For a given chemical process, consider the following equations:

$$\Delta G^* = \Delta H - T\Delta S^* = -RT\ln(K_{\text{eq}}) \quad (1.1.48)$$

$$\ln(K_{\text{eq}}) = -(\Delta H/R) T^{-1} + \text{constant} \quad (1.1.49)$$

where ΔG^* , ΔH and ΔS^* are the standard free energy, enthalpy and entropy changes, and K_{eq} is the equilibrium constant, or the mass-dependent part of the equilibrium conditions. For A–B molecular complexation, $K_{\text{eq}} = [\text{AB}]/[\text{A}][\text{B}]$, a measurable quantity. For evaporation and sublimation, the mass-dependent part is the saturated vapour pressure, so $K_{\text{eq}} = (\text{SVP})$, again a measurable quantity. Equation (1.1.49), the van't Hoff equation, allows an experimental determination of equilibrium enthalpies from equilibrium constants, for example, of sublimation

enthalpies of crystals from measurements of their SVP. The latter measurement is however a difficult one, because these pressures are of the order of 10^{-5} to 10^{-10} bar. Uncertainties on measured sublimation enthalpies are never smaller than 5–10 kJ mol⁻¹, and this should be kept in mind when using these numbers to calibrate intermolecular potentials.

Free energies can be simulated by the so-called free energy perturbation or thermodynamic integration methods. While, in principle, more revealing than enthalpy-based simulations, free energy simulations have a number of shortcomings. First, the sampling of phase space must be incomparably larger than in enthalpy simulations, so that the required calculations are computationally very demanding, and applications are limited to simple systems. Second, the choice of the computational parameters and of phase space sampling are critical for the significance and the numerical stability of the calculations. Finally, the theoretical underpinnings of free energy methods rely on subtle interpretations of statistical thermodynamics, so that great care must be applied when planning such calculations. Anyway, for reasons explained above, free energy differences between polymorphs cannot be too large, as they result from a sum of small ΔH s and small T ΔS contributions.

1.1.4.2 Melting and Polymorphism

Phase equilibria between condensed phases, like melting and crystal polymorphic transitions, have no mass-dependent terms (no equilibrium constants) since the activity of pure condensed phases is unity, and hence the equilibrium thermodynamics is represented by the simple relationships:

$$\Delta G = 0; \Delta S = \Delta H/T \quad (1.1.50)$$

Typical melting enthalpies of organic crystals are 10–50 kJ mol⁻¹, while the slope of a plot of experimental molar enthalpies of melting against melting temperature shows that the melting entropy of organic crystals is rather constant and is on average around 60 J K⁻¹ mol⁻¹. $\Delta H(\text{melt})$ can easily be simulated by MD by running separate simulations for the liquid and the crystal phase at the same temperature and taking the difference in total energies.

Enthalpy differences between polymorphic crystal forms of the same substance [18, 19] can be directly measured by calorimetry or by the difference in enthalpies of dissolution. A critical survey of these experiments shows that these enthalpy differences are very small, of the order of 0–10 kJ mol⁻¹, as expected since they must be a small fraction of the enthalpy of melting. In many cases the measured value is undistinguishable from experimental noise. These enthalpy differences can be estimated by just taking the differences in lattice energy between pairs of crystal phases whose complete structure is known, but the intrinsic uncertainty of such a calculation is also of the same order of magnitude of the property it tries to simulate.

1.1.5

Tutorial Examples

1.1.5.1 Dimerization Energies, a Scale of Intermolecular Interactions

Table 1.1.3 shows dimerization energies (Eq. (1.1.35), uncorrected for zero-point energy differences) of some sample bimolecular complexes, as calculated by Pixel, Eq. (1.1.34) [14]. Quite similar results are obtained by high-level *ab initio* calculations including electron correlation. These numbers constitute a relative scale of the strength of intermolecular interactions in organic compounds: the strongest is the hydrogen bond, with an energy of the order of 25–30 kJ mol⁻¹; next comes the arene–perfluoroarene stacking interactions, with about 20 kJ mol⁻¹ per ring, followed by aliphatic chain alignment, worth about 2 kJ mol⁻¹ per methylene group; next come the arene stacking interactions, whose strength per aromatic ring (around 10 kJ mol⁻¹) is about equal to some of the interactions between systems containing acidic C–H groups (e.g. the aromatics) and π -electron clouds of benzenoid rings. Of even lesser importance are most C–H \cdots O or C–H \cdots N type interactions, worth just a few kJ mol⁻¹, while other interactions (like Cl \cdots Cl contacts, often invoked as important in crystal structures) are uninfluential. Even more relevant to the discussion of recognition modes is the amplitude of the corresponding potential energy wells: Table 1.1.3 shows that only the hydrogen bond has a relatively narrow width, corresponding to a really competitive binding power. All other interactions merge into a highly stretchable continuum of energies and forces, so that their use for prediction and control of molecular recognition is very problematic and must be carefully justified in each instance [20].

Table 1.1.3 Equilibrium distances, binding energies and width of the potential energy well for hydrogen bonds.

Dimer	R ^[a]	ΔE ^[b]	Width ^[c]
benzoic acid, cyclic H-bond	1.75	67	1.65–1.95
formic acid, single H-bond	1.75	38	1.55–2.05
formamide, cyclic H-bond	2.0	52	1.85–2.2
formamide, single H-bond	2.05	25	1.8–2.45
methanol O–H \cdots O–H	1.95	23	1.7–2.3
formic acid–acetone O–H \cdots O=C	1.8	31	1.55–2.1
Water	2.0	23	1.85–2.2
parallel C ₆ H ₆ –C ₆ F ₆	3.5	23	3.25–4.0
parallel C ₆ F ₆ –C ₆ F ₆	3.55	17	3.28–4.0
parallel C ₆ H ₆ –C ₆ H ₆	3.70	10	3.4–4.4
parallel naphthalene	3.6	26	3.35–3.9
C–H \cdots π acetylene–acetylene	2.6	8	2.15–3.5
C–H \cdots π benzene–benzene ^f	2.65	11	2.25–3.5
CH ₃ \cdots π ethane–benzene ^f	3.2	7	2.75–4.3
contacts			
Acetone–acetylene C–H \cdots O	2.2	11	1.85–3.0

Table 1.1.3 (continued)

Dimer	R° ^[a]	ΔE ^[b]	Width ^[c]
Acetone–benzene C–H···O	2.45	6	2.05–4.0
Acetone–methane C–H···O	2.6	2.5	–
Acetonitrile–acetylene C–H···N	2.25	13	1.9–2.9
Acetonitrile–benzene C–H···N	2.5	6.5	2.15–4.2
Acetonitrile–methane C–H···N	2.6	2.5	–
Benzoquinone, two C=O···H–C parallel chain hexane	2.35	15	2.1–3.0
CH ₃ –Cl···Cl–CH ₃	–	0	–

[a] R° is the O···H or N···H distance in hydrogen bonds, or the inter-ring distance in parallel stacks, or the (C–)H··· π or (C–)H···O distance, Å units. [b] Binding energy, kJ mol⁻¹. [c] Range of the distance coordinate for an energy rise of 5 kJ mol⁻¹ from the minimum.

1.1.5.2 Calculation of Lattice Energies, Force Field Methods versus Pixel

The simplest crystal calculation consists of taking the atomic coordinates, cell data and space group for a crystal structure from an X-ray determination, building a static cluster of molecules to represent the crystal, and calculating all the interaction energies by some potential. This can be done by force field methods, Eq. (1.1.25) (time for one calculation a few milliseconds) or by Pixel, Eq. (1.1.34) (time for one calculation a few minutes). Figure 1.1.1 shows that the results are of comparable accuracy (remember that experimental values themselves are no more accurate than 5–10 kJ mol⁻¹). The UNI empirical force field [9] uses however about 100 parameters while the Pixel method uses only a few.

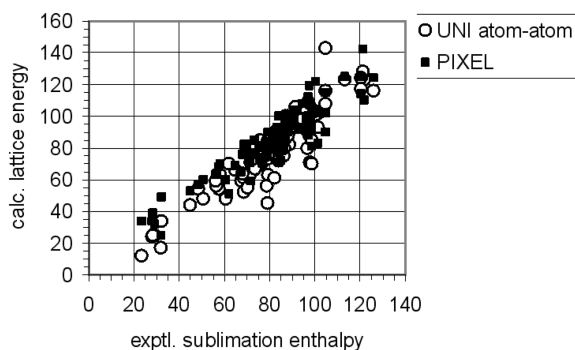


Figure 1.1.1 Experimental sublimation enthalpies and calculated lattice energies (kJ mol⁻¹) for 90 organic crystal structures.

1.1.5.3 Energy Partitioning by Pixel

Table 1.1.4 shows some typical examples of partitioned Pixel energies in molecular crystals. One sees there a clear predominance of dispersion in hydrocarbon crystals, an increased relevance of coulombic and polarization effects in heteroatom-containing molecules, and the clear predominance of coulombic-polarization energies in hydrogen-bonded crystals. In fact, a plot of the partitioned energies in hydrogen-bonded dimers shows that the total energy up to an $O \cdots H$ distance equal to the equilibrium value coincides almost exactly with the coulombic contribution. While of course the partitioning is to a large extent arbitrary and depends on the values of the parameters in the theory, the picture of intermolecular bonding is, within the present parametrization, quite convincing and helps explaining some of the main packing factors in crystals.

Table 1.1.4 Typical values for Pixel partitioned energies in some organic crystals, units kJ mol^{-1} .

Crystal name	E_{coul}	E_{pol}	E_{disp}	E_{rep}
hydrocarbons				
hexane	-10.7	- 3.8	-81.2	43.2
octane	-11.1	- 3.7	-95.5	45.9
benzene	-13.3	- 5.1	-59.5	30.1
anthracene	-26.1	-12.3	-115.9	68.4
1,4-benzene derivatives				
dichloro	-18.9	- 7.5	-86.2	45.6
dicyano	-42.3	-10.1	-92.6	51.2
dinitro	-39.3	-10.7	-86.4	37.6
benzoquinone	-36.0	- 8.6	-63.2	35.0
polar compounds				
succinic anhydride	-47.0	-12.3	-56.0	30.0
naphthoquinone	-36.7	-11.4	-105.9	51.8
triazine	-24.4	- 6.7	-68.1	34.9
acids				
acetic	-78.2	-35.8	-37.7	92.9
benzoic	-94.3	-46.2	-69.0	120.9
oxalic	-115.5	-47.7	-55.6	115.8
amides				
acetamide	-80.8	-27.7	-43.7	96.3
benzamide	-73.8	-25.9	-77.8	91.0
urea	-97.6	-32.3	-41.5	75.0

1.1.5.4 Analysis of Crystal Structures

The analysis of experimental crystal structures is at the same time the key to understanding the forces that determine one particular crystal structure, and a possible way to a better understanding of intermolecular forces in general. Methods based on geometrical features alone, like the identification of geometrical patterns, down to an analysis of atom–atom distances, are usually subjective and sometimes misleading. A comprehension of fine detail of the constitution of organic crystals requires objective, quantitative assessments. We propose the following systematic procedure for the quantitative classification of intermolecular interaction energies in crystals, consisting of the following sequential steps:

1. Using known atomic positions, cell dimensions and space group symmetry, find the closest molecular pairs in the crystal, for which the distance between centers of mass is smaller than the largest molecular dimension. The number of such neighbours is usually 10–12 in organic crystals.
2. Each molecular pair is characterized by the distance between centers of mass, R_i , a symmetry operator connecting the two molecules, O_i , and a molecule–molecule energy, E_i , calculated by some intermolecular potential method: quantum-chemical methods can be accurate, where affordable; the Pixel method has the advantage of good accuracy and energy partitioning; but even atom–atom force field energies can often be useful.
3. Rank these pairs in order of descending energetic relevance, and discuss the crystal structure in terms of these molecule–molecule determinants.
4. On the basis of these main interactions, proceed to the identification of extended motifs, like ribbons or layers within the crystal structure, where applicable.

We present here the analysis along these lines of some typical crystal structures [21]. Consider first naphthalene, a key example in aromatic interactions. Table 1.1.5 collects the interaction energies of closest neighbours, as calculated by the Pixel method, in comparison with the energies of some idealized, gas-phase dimers. In spite of many approximate and qualitative theories of aromatic packing, naphthalene in its crystal does not form close dimers, taking full advantage of the 50 kJ mol^{-1} dispersive power of parallel stacking at short interplanar distance (about 4 Å). Naphthalene molecules cannot form extended parallel stacks in the crystal because then lateral inter-stack C–H \cdots H–C contacts would not be stabilizing enough; the molecules are thus compelled to approach at high interplanar angles, using a moderate coulombic stabilization between electron-poor hydrogen regions and electron-rich π -clouds. In this perspective, the herring-bone structure of crystals of benzene and other aromatics is dictated by avoided repulsion, rather than by an often invoked C–H \cdots π attraction. Clearly, the observed crystal structure is a complex compromise between several demands, and does not lend itself to a simplification in terms of molecular geometry alone.

Table 1.1.5 Pixel energies (kJ mol^{-1}) between neighbor pairs in the naphthalene crystal and between optimized gas-phase dimers^[a] (PS and T).

Symmetry operator	Distance, Å.	E_{coul}	E_{pol}	E_{disp}	E_{rep}	E_{tot}
4 screw (A)	5.08	- 5.1	-2.9	-23.5	13.2	-18.3
2 translation (B)	5.97	- 4.5	-1.1	-14.0	6.4	-13.2
4 screw (C)	7.84	- 1.9	-0.9	- 7.7	5.0	- 5.6
PS	3.50	- 0.4	-5.0	-51.1	25.0	-31.6
T	4.88	-13.0	-7.9	-30.7	32.0	-19.6

[a] Idealized dimers: PS, parallel stacked, overlapping molecular planes; T, perpendicular molecular planes.

On the other hand, the most relevant molecular pairs in the crystal structure of naphthoquinone (Table 1.1.6) show a parallel stacked recognition mode. Molecules are related by a center of inversion, but contrary to naive electrostatic reasoning, the stacking arrangement is not stabilized by interaction between antiparallel dipoles, because the corresponding coulombic energies are very small. In contrast to naphthalene, naphthoquinone can use the parallel stacking mode because lateral inter-stack contacts are then stabilized by favorable coulombic adhesion between positively charged C–H moieties and oxygen basins (see determinants C, E, F in Table 1.1.6); note, however, that $\text{H}\cdots\text{O}$ short contacts do not always imply a sizeable coulombic interaction energy (see determinant D).

Table 1.1.6 Energies (kJ mol^{-1}) between neighbor pairs in the crystal of naphthoquinone.

Label	Distance	Short O \cdots H distances, Å	E_{coul}	E_{pol}	E_{disp}	E_{rep}	E_{tot}
1 A	4.07	none	- 2.4	-2.5	-45.7	13.9	-36.7
1 B	4.38	none	- 1.5	-1.7	-37.1	10.0	-30.3
1 C	8.13	2.52	-14.8	-3.3	- 8.6	11.6	-15.0
2 D	6.65	2.87, 2.71	- 2.2	-1.6	-12.8	5.8	-10.9
1 E, 1 F	7.40	2.47, 2.49	-13.2	-3.6	-10.2	14.9	-12.0

In the naphthoic acid crystal (Table 1.1.7), the by far most dominant structure determinant is, as expected, the hydrogen bonded cyclic dimer, A, described by the Pixel approach as a strongly coulombic interaction, while the polarization component supposedly represents the stabilization deriving from the dynamic rearrangement of the electron density – the partial covalent bond character. Next, at a considerable energetic distance, comes a determinant in which the dispersive stacking contribution dominates. The hydrogen bonded dimer is complemented

by two in-plane molecular interactions (E–F), using the favorable C–H···O coulombic pairing as in naphthoquinone. Whether one would like to call these interactions C–H···O bonds, is a matter of taste; energy-wise, this may not be advisable, because even the sum of the REF-E and REF-F energies, $-14.4 \text{ kJ mol}^{-1}$, is still smaller than the smaller of the energies in the stacked dimers.

Table 1.1.7 Energies (kJ mol^{-1}) between neighbour pairs in the crystal of 2-naphthoic acid.

Label	Distance, Å	E_{coul}	E_{pol}	E_{disp}	E_{rep}	E_{tot}
1 A	9.68	-146.0	-78.2	-13.5	183.8	-53.8
2 B	5.00	-6.0	-3.1	-37.6	19.6	-27.1
1 E	8.29	-7.9	-3.0	-5.4	5.8	-10.4
2 F	7.53	-1.9	-1.7	-7.5	7.1	-4.0

References

- 1 Hehre, W. J., Radom, L., Schleyer, P. v. R., Pople, J. A., *Ab Initio Molecular Orbital Theory*, Wiley, New York 1986.
- 2 Ziegler, T. Approximate density functional theory as a practical tool in molecular energetics and dynamics, *Chem. Rev.* **1991**, *91*, 651.
- 3 Bickelhaupt, F. M., Baerends, E. J., Kohn-Sham density functional theory: predicting and understanding chemistry, *Rev. Comput. Chem.*, **2000**, *15*, 14.
- 4 Grimme, S., Accurate description of van der Waals complexes by density functional theory including empirical corrections, *J. Comput. Chem.* **2004**, *25*, 1463.
- 5 Dovesi, R., Civalieri, B., Orlando, R., Roetti, C., Saunders, V. R., Ab initio quantum simulations in solid state chemistry, *Rev. Comput. Chem.* **2005**, *21*, 1.
- 6 Williams, D. E., Starr, T. L., Calculation of the crystal structures of hydrocarbons by molecular packing analysis, *Comput. Chem.* **1977**, *1*, 173.
- 7 Kitaigorodski, A. I., in *Advances in Structure Research by Diffraction Methods*, Brill, R., Mason, R. (Eds.), Pergamon Press, Oxford 1970, Vol. 3, p.173.
- 8 G. Filippini, G., Gavezzotti, A., Empirical intermolecular potentials for organic crystals: the 6-exp approximation revisited, *Acta Crystallogr. Sect. B* **1993**, *49*, 868.
- 9 Gavezzotti, A., Filippini, G., Geometry of the intermolecular X–H···Y (X, Y = N, O) hydrogen bond and the calibration of empirical hydrogen-bond potentials, *J. Phys. Chem.* **1994**, *98*, 4831.
- 10 Williams, D. E., Net atomic charge and multipole models for the ab initio molecular electric potential, in *Rev. Comput. Chem.* **1991**, *2*, 219.
- 11 van Gunsteren, W. F., Berendsen, H. J. C., Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992.
- 12 Gavezzotti, A., Calculation of intermolecular interaction energies by direct numerical integration over electron densities. I. Electrostatic and polarization energies in molecular crystals, *J. Phys. Chem.* **2002**, *106*, 4145.
- 13 Gavezzotti, A., Calculation of intermolecular interaction energies by direct numerical integration over electron

- densities. 2. An improved polarization model and the evaluation of dispersion and repulsion energies, *J. Phys. Chem. B*, **2003**, *107*, 2344.
- 14** Gavezzotti, A., Quantitative ranking of crystal packing modes by systematic calculations on potential energies and vibrational amplitudes of molecular dimers, *J. Chem. Theor. Comput.* **2005**, *1*, 834.
- 15** Pertsin, A. J., Kitaigorodski, A. I., *The Atom-Atom Potential Method, Applications to Organic Molecular Solids*, Springer-Verlag, Berlin 1987, Ch. 5 and 6.
- 16** Kollmann, P., Free energy calculations: applications to chemical and biochemical phenomena, *Chem. Rev.* **1993**, *93*, 2395.
- 17** Reinhardt, W. P., Miller, M. A., Amon, L. M., Why is it so difficult to simulate entropies, free energies, and their differences? *Acc. Chem. Res.* **2001**, *34*, 607.
- 18** Gavezzotti, A., Filippini, G., Polymorphic forms of organic crystals at room conditions: thermodynamic and structural implications, *J. Am. Chem. Soc.* **1995**, *117*, 12299.
- 19** Bernstein, J. *Polymorphism in Organic Crystals*, Oxford University Press, Oxford 2002.
- 20** Dunitz, J. D., Gavezzotti, A., Molecular recognition in organic crystals: directed intermolecular bonds or nonlocalized bonding? *Angew. Chem. Int. Ed.* **2005**, *44*, 1766.
- 21** Gavezzotti, A., Hierarchies of intermolecular potentials and forces: progress towards a quantitative evaluation, *Struct. Chem.* **2005**, *16*, 177.