Molecular Self-Assemblies. 5. Analysis of the Vector Properties of Hydrogen Bonding in Crystal Engineering

Jerry Perlstein,*,† Kevin Steppe,† Shai Vaday,† and Edmund Moses N. Ndip‡

Contribution from the Center for Photoinduced Charge Transfer, Department of Chemistry, The University of Rochester, Rochester, New York 14627, and Department of Chemistry, Hampton University, Hampton, Virginia 23668

Received December 29, 1995. Revised Manuscript Received May 6, 1996[⊗]

Abstract: Kitaigorodskii's aufbau principle (KAP) is used to analyze hydrogen bonding as a vector for the packing of molecules in the crystalline solid state. Using the CFF91 force field to compute the molecule−molecule interaction potential, we find that the signature for hydrogen bonding of N−H and O−H donor groups with N or O acceptors is a positive value for the nonbonded van der Waals term of the hydrogen atom involved in the H-bond. The H-bond may occur as a vector in any one or more of the four stages of KAP. We categorize these vectors as types 1−4 with 16 possible subtypes depending upon the number of KAP stages in which the H-bond appears. Within the constraints of the force field description, the H-bond then becomes a specific vector contribution to the packing of one or more substructures of the complete crystal. Of the 16 possible vector subtypes we illustrate 12 of them using crystal structure data from the Cambridge Structural Database. Knowing the vector subtype, we show how it is possible to locate the local minima representing the packing geometry of the substructures using Monte Carlo simulation methods. The establishment of quantitative signatures for vector types within KAP substructures and their relationship to crystal engineering is discussed.

Introduction

Supramolecular chemistry, molecular self-assembly, molecular recognition, crystal engineering, and other similar phrases are used as descriptors of molecule−molecule interactions which involve noncovalent or van der Waals "bonds". The packing of molecules into arrays of varying and beautiful structures is the result of this noncovalent interaction of which crystals with regular repeating molecular units are one example.[1] One of the goals of the supramolecular chemist is to be able to design such arrays a priori knowing only the stereochemistry of the atom−atom connectivity of the molecules. Unfortunately, very little is known about the "stereochemistry" of the noncovalent bond. The closest equivalent to the covalent bond for supramolecular engineering of noncovalent interactions is the hydrogen bond, whose quantum chemical description is still very much under discussion.[2] Nevertheless, the H-bond has been used with considerable success as a stereochemical force or vector in qualitatively designing molecular arrays of various shapes.[3] The various geometric patterns which the H-bond can assume in these structures has been categorized in terms of graph sets by Etter,[4] which has been most useful for searching a structure for H-bond interactions. Aakeröy and Seddon[1e] call the hydrogen bond in such cases a synthetic "vector" since it acts as a regiospecific structure controlling agent. We will adopt their phrase and refer to hydrogen bonds as H-bond vectors to describe this type of control throughout this paper.

Chemists who have used the H-bond as a design tool have been very clever in placing the H-bond donors and acceptors in just the right places on the molecular frame to achieve desirable results.[3] The H-bonds in such cases control the interaction to such an extent as to be the dominant force for the packing of the molecules containing them. Below we will show how this comes about, but this situation is not generally the case. Given an arbitrarily shaped organic molecule with H-bond donors and acceptors, the packing geometry is not so nearly discernible a priori even though H-bonds may be abundantly present. The H-bond is only one vector in a multitude of other noncovalent vectors with which it competes. For example, in the gas and solution phases acetic acid forms the well-known inversion dimer shown schematically in Scheme 1, but in the solid state the structure is the glide chain also shown in Scheme 1.[5] Clearly there are other vectors involved here.[6]

It is the purpose of this paper to put the vector concept of the hydrogen bond into a more general perspective, to show how and when it expresses itself as a vector to promote a specific

* To whom all correspondence should be addressed. E-mail address: perlstein@chem.chem.rochester.edu.

† The University of Rochester.

‡ Hampton University.

⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1996.

(1) For leading references, see the special issue dedicated to Margaret Etter, "Structure and Chemistry of the Organic Solid State": *Chem. Mater.* **1994**, *6*, 1087−1461. See also the following reviews: (a) MacDonald, J. C.; Whitesides, G. M. *Chem. Rev.* **1994**, *94*, 2383−2420. (b) Webb, T. H.; Wilcox, C. S. *Chem. Soc. Rev.* **1993**, 383−395. (c) Zaworotko, M. J. *Chem. Soc. Rev.* **1994**, 283−288. (d) Subramanian, S.; Zaworotko, M. J. *Coord. Chem. Rev.* **1994**, *137*, 357. (e) Aakeröy, C. B.; Seddon, K. R. *Chem. Soc. Rev.* **1993**, 397−407.

(2) (a) *Modeling the Hydrogen Bond*; Smith, D. A., Ed.; ACS Symposium Series 569; American Chemical Society: Washington, DC, 1994. (b) Jurema, M. W.; Shields, G. C. *J. Comput. Chem.* **1993**, *14*, 89−104. (c) Dannenberg, J. J.; Evleth, E. M. *Int. J. Quantum Chem.* **1992**, *44*, 869−885. (d) Scheiner, S. In *Reviews in Computational Chemistry II*; Lipkowitz, K. B., Boyd, D. B., Ed.; VCH: New York, 1991; pp 165−218. (e) Dykstra, C. E. *Acc. Chem. Res.* **1988**, *10*, 355−361.

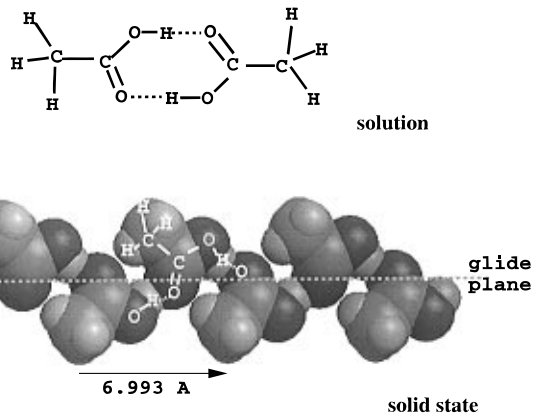
(3) (a) Subramanian, S.; Zaworotko, M. J. *Can. J. Chem.* **1995**, *73*, 414. (b) Hanessian, S.; Simard, M.; Roelens, S. *J. Am. Chem. Soc.* **1995**, *117*, 7630−7645. (c) Kane, J. J.; Liao, R.-F.; Lauher, J. W.; Fowler, F. W. *J. Am. Chem. Soc.* **1995**, *117*, 12003−12004. (d) Russell, V. A.; Etter, M. C.; Ward, M. D. *J. Am. Chem. Soc.* **1994**, *116*, 1941−1952. (e) Zerkowski, J. A.; MacDonald, J. C.; Seto, C. T.; Wierda, D. A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 2382−2391. (f) Chang, Y.-L.; West, M.-A.; Fowler, F. W.; Lauher, J. W. *J. Am. Chem. Soc.* **1993**, *115*, 5991−6000. (g) Yank, J.; Fan, E.; Geib, S. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **1993**, *115*, 5214−5315. (h) Etter, M. C.; Reutzel, S. M.; Choo, C. G. *J. Am. Chem. Soc.* **1993**, *115*, 4411−4412. (i) Pascal, R. A., Jr.; Ho, D. M. *J. Am. Chem. Soc.* **1993**, *115*, 8507−8508. (j) Wyler, R.; Mendoza, J. D.; Rebek, J. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1699−1701. (k) Simard, M.; Su, D.; Wuest, J. D. J. *J. Am. Chem. Soc.* **1991**, *113*, 4696−4698.

(4) (a) Etter, M. C. *J. Phys. Chem.* **1991**, *95*, 4601−4610. (b) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120−126. (c) Bernstein, J.; Davis, R. E.; Shimoni, L.; Chang, N.-L. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1555−1573.

(5) Jonsson, P.-G. *Acta Crystallogr., Sect. B* **1971**, *27*, 893.

S0002-7863(95)04341-1 CCC: $12.00

Scheme 1



packing pattern in the crystalline solid state and how and when other noncovalent forces compete with it. The present work is an outgrowth of our previous studies based on Kitaigorodskii's aufbau principle (KAP), which we have described in detail,[7] coupled with a force field description of the hydrogen bond and other nonbonded interactions. The force field we use here is CFF91 from the force field consortium of Biosym/MSI.[8]

Kitaigorodskii's Close Packing and Etter's Rules for Hydrogen Bond Formation

In the late 1950s and early 1960s Kitaigorodskii gave a geometric description of crystal structures based on close-packing principles.[9] He described the most likely packing arrangements of arbitrary shaped organic molecules assuming that Nature would like them to be close packed, viz., filling as much of space as possible leaving a minimum of void density. In contrast to this, we have Etter's first rule,[4] namely, "all good proton donors and acceptors are used in hydrogen bonding." While Etter's rule would appear at first blush to be contrary to Kitaigorodskii's close-packing arguments, it has nevertheless been shown by Gavezzotti that close packing is obeyed even for the vast majority of organic molecules that make hydrogen bonds in the solid state.[10] It is most remarkable that arbitrary shaped molecules containing hydrogen bond donors and acceptors fill space so completely and yet at the same time make as many hydrogen bonds as possible consistent with the donor−acceptor ratio. How is it possible for Kitaigorodskii's space-filling feature of solids and Etter's hydrogen bond making feature of solids to be simultaneously satisfied? Crystal engineers have tended to focus on H-bond vectors as the controlling forces for crystal packing because of their relative strength and directionality. One can also visualize the hydrogen bond in a line drawing, which is not so easily done for close packing. In order to put these two ideas into better perspective, we will analyze these two concepts in this paper and show that close packing and hydrogen bonding come together in a natural way. In particular we wish to provide a framework for hydrogen bonding within the context of Kitaigorodskii's packing arguments with the intent of providing a simple yet satisfactory way to (a) analyze hydrogen-bonded crystals and (b) provide a general framework in which the vector concept for crystal engineering may be exploited.

Kitaigorodskii's Aufbau Principle

Kitaigorodskii described the packing of molecules in crystals in terms of close-packed substructures[9] which we have recently shown quantitatively to be local energy minima (or at worst stationary points) of the molecular interaction potential.[11] Using Monte Carlo annealing methods[11] or conformational analysis methods,[12] these substructures can be shown to be quantitatively predictable. What are these substructures? Kitaigorodskii described three types; we include a fourth for completeness. We refer to them as stages in the aufbau of the complete crystal, and they represent the lowest energy substructures of the crystal. These stages and how to compute their energy are described below.

A. Stage 0. This stage defines a molecular "packing unit" which is used to compute molecule−molecule interaction energies. The packing unit is the unit relative to which the lattice energy is computed. In contrast to the crystallographic asymmetric unit, a packing unit is the smallest number of *whole* molecules with which the crystal structure can be constructed. Fortunately, for most crystal structures (more than 90% of them) containing one molecule in the crystallographic asymmetric unit, the packing unit and the asymmetric unit are identical. They are also identical if the asymmetric unit consists of two or more *whole* molecules (but these have to be chosen carefully so that the interaction potential between the molecules is the lowest possible). However, if the asymmetric unit is only half a molecule (for example, if it sits on an inversion center), the packing unit is still one whole molecule simply found by completing the molecule with the inversion-related atoms prior to packing it. More complex asymmetric units with partial molecules can exist. In each case, the partial molecules should be completed so as to define a packing unit with an integral number of whole molecules and whose intermolecular interaction energy is the lowest possible. This stage can get very complex. But regardless of its complexity, it has one important feature, namely, that it consists of a finite number of molecules as contrasted with the remaining stages, which contain an infinite number of molecules. It can be a small molecule with intramolecular H-bonds or a large complex molecular structure like a protein. It may also consist of several component molecules such as a heterodimer or a complex structure like a phospholipid vesicle. In this paper we will confine ourselves to simple stage 0 structures in which the packing unit is one whole molecule. The arguments presented here should be extendible to more complex stage 0 arrays.

B. Stage 1. This stage is determined using the packing unit of stage 0 as illustrated by the acetic acid glide chain in Scheme 1. The packing unit is a single acetic acid molecule, which is then packed into a one-dimensional infinite chain, which has glide symmetry in this case. It has the following characteristics:

(1) The chain is infinite with a single repeat vector.

(6) For example, weaker C−H···O bond vectors have been invoked which we do not consider in this paper. For discussions on this type of interaction, see the following classic papers: (a) Berkovitch-Yellin, Z.; Leiserowitz, L. *J. Am. Chem. Soc.* **1982**, *104*, 4052−4064. (b) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063−5070. For a review, see: Desiraju, G. R. *Acc. Chem. Res.* **1991**, *24*, 290−296. For quantum chemical treatments, see, for example: (a) Turi, L.; Dannenberg, J. J. *J. Am. Chem. Soc.* **1994**, *116*, 8714−8721. (b) Koch, U.; Popelier, P. L. A. *J. Phys. Chem.* **1995**, *99*, 9747−9754.

(7) Perlstein, J. *J. Am. Chem. Soc.* **1994**, *116*, 11420−11432.

(8) (a) Hwang, M. J.; Stockfisch, T. P.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2515−2525. (b) Maple, J. R.; Hwang, M.-J.; Stockfisch, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *J. Comput. Chem.* **1994**, *15*, 162−182.

(9) Kitaigorodskii, A. I. *Organic Chemical Crystallography*; Consultants Bureau: New York, 1961; Chapter 3, pp 65−112.

(10) Gavezzotti, A.; Filippini, G. *J. Phys. Chem.* **1994**, *98*, 4831−4837.

(11) (a) Perlstein, J. *Chem. Mater.* **1994**, *6*, 319−326. (b) Perlstein, J. *J. Am. Chem. Soc.* **1994**, *116*, 455−470. (c) Perlstein, J. *J. Am. Chem. Soc.* **1992**, *114*, 1955−1963.

(12) (a) Scaringe, R. P. In *Electron Crystallography of Organic Molecules*; Fryer, J. R., Dorset, D. L., Eds.; Dordrecht, 1990; pp 85−113. (b) Scaringe, R. P. *Trans. Am. Crystallogr. Assoc.* **1992**, *28*, 11−21.

(2) The chain repeat need not occur along any X-ray-derived unit-cell axis. (For acetic acid the lowest energy chain occurs along a face diagonal.)

(3) Except for a simple translation chain, the centroids of the molecules in the chain do not lie on the chain axis, but may be offset from it by equal distances. This can make the chain appear to be several molecules thick.

There are 75 ways to pack a stage 0 cluster in one dimension.[13] Most crystals crystallize in space groups with orthorhombic or lower symmetry.[14] As shown by Kitaigorodskii[9] and beautifully elaborated by Scaringe and Perez,[15] with the unlikelihood of simple 2-fold rotation axes or mirror plane symmetry, only four of these chains are of any statistical significance. These are translation, 2_1 screw, glide, and inversion chains. Details of the packing patterns for these chains have been provided elsewhere.[11]

C. Stage 2. Stage 2 is a continuation of the aufbau from one dimension into two. It is determined from stage 1 chains by repeating them in a second dimension with appropriate symmetry. There are 85 possible layers of this type elaborated by Wood,[16] but as shown by Scaringe,[12a] only seven of these are of any statistical significance for crystals with orthorhombic or lower symmetry. All stage 2 structures have the following common characteristics:

(1) They are infinite in extent, having two repeating vectors. The repeat vectors for the lowest energy layer need not occur along the X-ray-derived unit-cell vectors of the crystal.

(2) Like stage 1, the molecules usually do not have their centroids all lying in the layer plane. This results in layers which (i) have coarse rather than smooth surfaces and (ii) can be more than one molecule thick.

D. Stage 3. The final stage is the full three-dimensional crystal packing determined by combining the stage 2 layers with appropriate symmetry in the third dimension.

There is nothing in the above description that says anything about hydrogen bonding. Only energetic arguments are used to determine the composition of each stage. Since most H-bonded crystals in stage 3 are close-packed, hydrogen bonding is not a special situation for KAP. Hydrogen-bonded crystals can therefore be described by the same four stages of KAP as for any other crystal. What we propose to show by examples below is that hydrogen bonding as a vector is a natural extension of KAP. It occurs as a contributory factor but not the only factor in the packing energy of each stage. To do this in a quantitative way we need an energetic description of the packing which will allow the determination of the geometry for each KAP stage as well as an evaluation of the presence of hydrogen bonds within those stages. We also need a general way of descriptively characterizing the results so that the H-bond vector characteristics can be summarized for any crystal structure.

Crystal Packing Energetics

We use an atom−atom potential force field approach for determining the energetic properties of each KAP stage. Various force fields can be used for this purpose. Our choice was based on ease of implementation and expected quality of results. The CFF91 force field from Biosym/MSI suited our purposes nicely since it contains no special substructures and no special hydrogen-bonding equations.[17] It also is able to quantitatively locate hydrogen-bonded structures as local energy minima, which we show below. Only two terms are needed from this force field to adequately describe the molecule−molecule interaction potential including hydrogen bonding as follows:

$$E^{\mathrm{nb}} = \sum_{i,j} A_{ij}\left[2\left(\frac{B_{ij}}{r_{ij}}\right)^9 - 3\left(\frac{B_{ij}}{r_{ij}}\right)^6\right] \tag{1}$$

$$E^{\mathrm{el}} = \sum_{i,j} \frac{q_i q_j}{\epsilon_0 r_{ij}} \tag{2}$$

$$E^{\mathrm{lattice}} = \frac{1}{2}(E^{\mathrm{nb}} + E^{\mathrm{el}}) \tag{3}$$

E^{nb} in eq 1 is the sum of the nonbonded van der Waals (Vdw) dispersive and repulsive terms with parameters A_{ij} and B_{ij} given by the force field, and r_{ij} is the atom−atom distance. The r^{-9} dependence of the repulsive part of this potential is considerably softer than the r^{-12} dependence seen in the Lennard-Jones potential. In eq 2, E^{el} is a Coulomb electrostatic term where we use a dielectric constant $\epsilon_0 = 1.0$. The summations in eqs 1 and 2 are over all atoms i of a stage 0 packing unit in the center of the close-packed crystal interacting with all atoms, j, in all the other molecules of the crystal. The lattice energy, E^{lattice}, for each stage of KAP is then computed from eq 3 with the $^1/_2$ inserted to correct for double counting. It is important to point out here that the partial atomic charges, q, in eq 2 are those given by the force field and not determined by some other computational or parametrization scheme. Parameters A and B and incremental charges for computing the q's have been published by Biosym/MSI[17] and are included here as supporting information. There are also no specific H-bonded terms in this force field. The entire parametrization is controlled by the atom types. Once the atom types are set, the atom−atom energy terms are readily computed. Missing from eq 2 is any polarization correction to the stage 3 energy for those crystals which have a dipole moment.[18] (This correction is not important for the other stages.) To include this effectively requires doing an Ewald sum of the electrostatic potential rather than the direct sum.[19] However, since our concern here is with the occurrence of hydrogen bonding rather than a complete energetic analysis, we have omitted this from the computations.

Identifying the Packing Geometry of Each Stage in Known Crystal Structures

To determine the molecules involved in each stage of KAP assume the following for the moment that the stage 0 packing unit consists of one molecule:

(1) Pack the molecule in the space group with sufficient unit cells so that the packing unit is completely surrounded by other molecules.

(2) Compute the interaction between the packing unit and each and every other molecule around it using eq 3.[20]

(3) Find the lowest energy interaction. The symmetry of this interaction is the chain type for stage 1. If the symmetry is translation, glide, or 2_1 screw, there will always be two identical

(13) Shubnikov, A. V.; Koptsik, V. A. *Symmetry in Science and Art*; Plenum: New York, 1974; p 114.
(14) Brock, P. B.; Dunitz, J. D. *Chem. Mater.* **1994**, *6*, 1118−1127.
(15) Scaringe, R. P.; Perez, S. *J. Phys. Chem.* **1987**, *91*, 2394−2403.
(16) Wood, E. A. *Bell Syst. Tech. J.* **1964**, *43*, 541−559 and references therein to the earlier categorization of the layer groups.
(17) *Discover User Guide, Version 2.8 Part 2*; Biosym Technologies: San Diego, March 1992. A reasonably complete set of parameters for the CFF91 force field are listed in this version.
(18) Smith, E. R. *Proc. R. Soc. London, A* **1981**, *375*, 475−505.
(19) For a nice example which demonstrates the difference between the direct lattice sum and the Ewald lattice sum as applied to the polymorphs of glycine, see: Derissen, J. L.; Smit, P. H.; Voogd, J. *J. Phys. Chem.* **1977**, *81*, 1474−1476.

lowest molecule−molecule energy interactions which completely define the stage 1 structure. If the lowest energy interaction symmetry is inversion, find the second-lowest energy interaction with inversion symmetry. This will then define an inversion chain for stage 1. We found it helpful in doing this to color the molecules in the crystal according to symmetry for ease of identification as described previously.[11b] The packing unit and all molecules related to it by simple translation are given one color. All molecules related to it by other symmetry elements are given a different color (one color for each symmetry element). CHEMX provides a convenient facility for carrying this out.[21] As shown in the various figures, the molecules which make up a chain generally lie with their centroids not on the chain axis (except for the translation chain) but offset from the chain axis.

(4) Once a stage 1 chain has been found, the stage 2 layer is easy to find as follows: Looking down the repeat vector of the stage 1 chain, one sees this chain surrounded by other chains. Compute the interaction potential of the packing unit with the molecules in each of these surrounding chains. If the lowest packing unit−chain interaction is related to the stage 1 chain by simple translation, then you are done. The two chains now form a stage 2 layer with two repeat vectors. This is the simplest form of layer formation. If the lowest packing unit chain interaction is not related to the stage 1 chain by simple translation, then look for a second surrounding chain with the same type of interaction. This pair will then provide the second repeat vector for stage 2. It is worthwhile to sketch a little diagram showing the chain containing the packing unit and the surrounding chains with the energy marked down for each packing unit−chain interaction. In this way the energy for pairs of chains can be computed to find the lowest energy layer.

(5) Stage 3 is the whole crystal. The energy for this stage is simply that for the packing unit interacting with all the molecules summed according to eq 3. We again note that in a more exact treatment for polar space groups this energy should be corrected for the polarization energy.

Multiple Molecules in Stage 0

If stage 0 contains more than one molecule, say two different molecules, the procedure is the same as for one molecule except now the packing unit is both molecules taken together. Before doing this, however, it is important to establish which pair of molecules in the crystal structure has the lowest interaction energy. This usually will not be the two molecules in the unit cell as given by the crystallographer, but is easy enough to determine by simply packing the crystal and computing the interaction potential between any one molecule and all the others. The pair that has the lowest energy represents stage 0. This pair should be used as the packing unit to repack the crystal. Again, coloring the molecules according to symmetry is very helpful in sorting out the symmetry relations of this more complex situation.

The Signature for Hydrogen Bonding

Equation 1 also contains the signature for intermolecular hydrogen bonds. We have found the following:

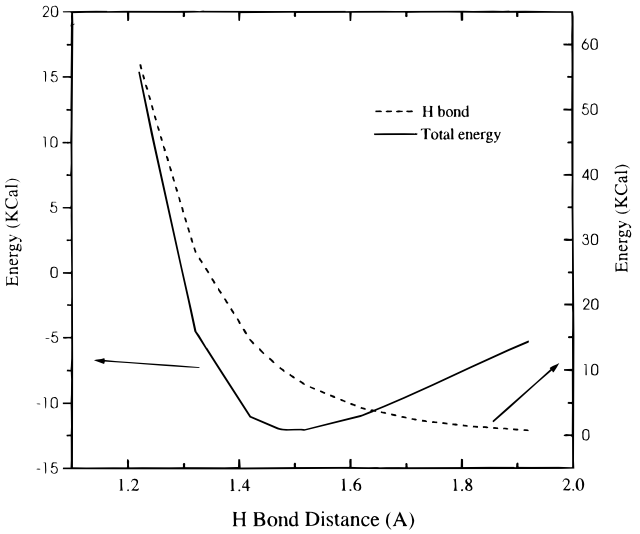


Figure 1. Energy vs donor H−acceptor O separation distance: (− − −) nonbonded van der Waals energy as computed from eq 1; (—) total intermolecular energy as computed from eq 3.

When an intermolecular hydrogen bond is present in a structure containing O or N hydrogen donors, the contribution to eq 1 of the donor hydrogen atom interacting with O or N acceptors is always repulsive (positive sign). In the absence of a hydrogen bond, the hydrogen atom contribution is either negative or zero.

Hydrogen atom position between the donor and acceptor can be quite variable as indicated by Jeffrey and Saenger.[22] We therefore tested the above assertion by computing the H-bond contribution to the total interaction energy of two H-bonded inversion-related carboxylic acid molecules, keeping the donors and acceptors spatially fixed at 2.62 Å but allowing the hydrogen atom to move from a long distance from the acceptor (1.92 Å) to a short distance (1.22 Å), a distance which covers the vast majority of H-bond structures. The results shown in Figure 1 indicate that, no matter where the hydrogen is placed, its contribution to the van der Waals energy (eq 1) is always positive while the total intermolecular energy has a minimum. In practice this means that donor hydrogens that are missing from a molecule's X-ray crystal structure can simply be added at a conventional distance with the knowledge that the H-bond signature will be fulfilled. However, it also means that one should not put too much faith in the absolute value of the H-bond nonbonded Vdw contribution as any measure of H-bond "strength".

Unfortunately, our attempt to find a signature for *intramolecular* H-bonding was not successful. Because of the generally close approach of atoms in a covalently bonded structure, the donor H atom Vdw term always turned out to be positive whether an intramolecular H-bond was present or not. We therefore relied on the visual overlap of the van der Waals spheres for the donor hydrogen with the acceptor to determine the presence of an intramolecular stage 0 H-bond. Further work needs to be done in this area to find an appropriate intramolecular H-bond signature.

Cataloguing the H-Bonds

Since there are four stages of KAP and hydrogen bonds can occur in any one or more of them, the total number of such possibilities is 16 (including no H-bond at all) as detailed in Table 1. We use the following notation to describe these

(20) Fortran routines for computing this energy as interfaces to the CHEMX/CHEMLIB molecular modeling package are available from the authors upon request. Contact the authors at internet address perlstein@chem.chem.rochester.edu.

(21) CHEMX is a molecular modeling program developed and distributed by Chemical Design Ltd., Roundway House, Cromwell Park, Chipping Norton, Oxfordshire OX7 5SR, U.K. CHEMLIB is a CHEMX interface which allows the user to link subroutines to CHEMX.

(22) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1991; p 95.

Table 1. The 16 Possible Hydrogen-Bonded Vector Types and Subtypes in the Four Stages of Kitaigorodskii's Aufbau

type no.	KAP stages	subtype notation	H-bond descriptor
0		0	No Hydrogen-bonds present
1			H-bond vector present in only one stage
	0	1_0	intramolecular H-bond (but can also be intermolecular for a finite group of molecules)
	1	1_1	H-bonds occur in a chain structure (intrachain) with typically translation, screw, glide, or inversion symmetry
	2	1_2	a vector which couples chains from stage 1 (interchain) to make a layer usually one of seven layer types
	3	1_3	H-bond occurs only between layers (interlayer vector)
2			H-bond vector present in only two stages
	0, 1	2_{01}	intramolecular vector and intrachain vector occurring simultaneously
	0, 2	2_{02}	intramolecular vector and interchain vectors making layers
	0, 3	2_{03}	intramolecular vector and interlayer vectors only
	1, 2	2_{12}	only intrachain and intralayer vectors present
	1, 3	2_{13}	only intrachain and interlayer vectors present
	2, 3	2_{23}	only intra- and interlayer vectors present
3			H-bond vectors occurring in three stages
	0, 1, 2	3_{012}	intramolecular H-bonds as well as intrachain and intralayer vectors
	0, 1, 3	3_{013}	intramolecular H-bonds as well as intrachain and interlayer vectors
	0, 2, 3	3_{023}	intramolecular H-bonds as well as intra- and interlayer vectors
	1, 2, 3	3_{123}	no intramolecular vectors but H-bonds in all other stages
4	0, 1, 2, 3	4_{0123}	H-bond vectors occurring in all four stages

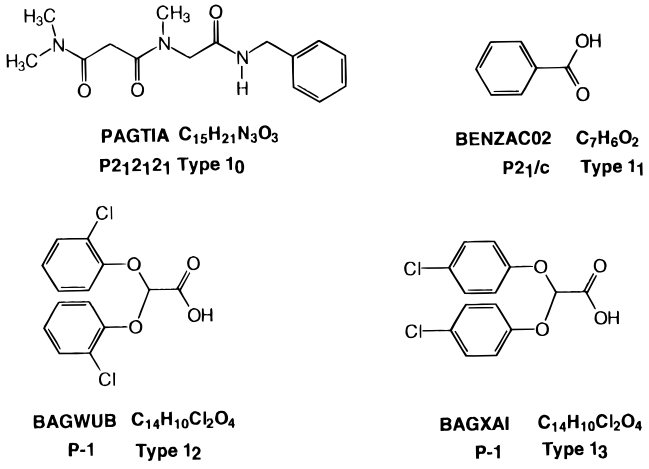


Figure 2. Molecular structures, formulas, space groups, and subtypes belonging to type 1 hydrogen-bonded crystals. Refcodes are from the Cambridge Structural Database.[24]

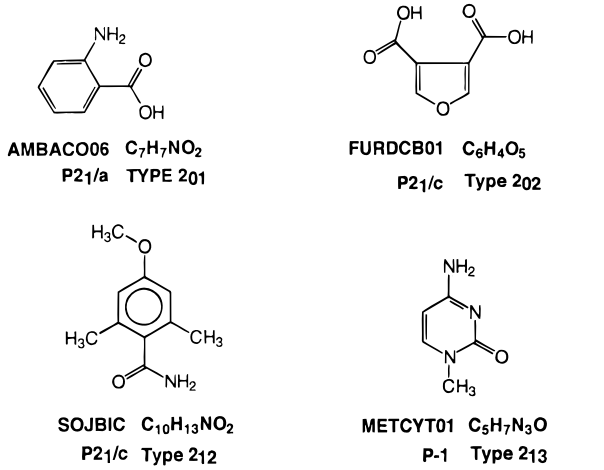


Figure 3. Molecular structures, formulas, space groups, and subtypes belonging to type 2 hydrogen-bonded crystals. Not shown are types 2_{03} and 2_{23}. Refcodes are from the Cambridge Structural Database.[26]

H-bond types. Type 1 means that a hydrogen bond occurs in one and only one stage of KAP. Type 2 means that it occurs in only two stages and so on. We also denote by subscript the specific stages where an H-bond occurs within the given type. Thus type 2_{13} means that H-bonds occur in two stages, namely, stage 1 and stage 3 and no others.

Results

In Figures 2−4 we present examples for most of the H-bond types in crystals whose structures are taken from the Cambridge Structural Database.[23] We will refer to these by their refcode names. To simplify the discussion, we have limited our choice of examples to those for which stage 0 consists of only a single molecule. Extension to more complex situations is reasonably straightforward. In Tables 2−4 are the energy terms in kilocalories/mole for each stage of KAP, except stage 0, as computed using eqs 1−3. Column 2 is the Vdw energy contribution to eq 3 excluding the donor hydrogen atom−acceptor interaction; column 3 contains the specific donor H atom−acceptor contribution indicative of the signature for the H-bond vector; column 4 is the total electrostatic contribution to eq 3; and column 5 is the total lattice energy (the sum of the terms in columns 2−4).

A. Type 0: No H-Bond Vectors. This is the null case and has been discussed in some detail previously.[7,11] The packing of molecules in the various stages of KAP is determined mainly by the nonbonded Vdw term. Monte Carlo[7,11] and conformational methods[12,15] for finding the energy minima for these stages have been described. Even when donor hydrogens are present, E^{nb}(H-bond) is less than or equal to 0, in violation of Etter's rule. Hydrogen bonding cannot compete with the other Vdw energy terms for these cases.

B. Type 1 H-Bond Vectors. Figure 2 shows examples of molecules with type 1 H-bond vectors.[24] All these structures

(23) (a) Allen, F. H.; Kennard, O. *Chem. Design Automat. News* **1993**, *8*, 31−37. (b) Allen, F. H.; Kennard, O.; Taylor, R. *Acc. Chem. Res.* **1983**, *16*, 146−153.

(24) Refcodes are from the Cambridge Structural Database and are as follows. (a) PAGTIA: Dado, G. P.; Desper, J. M.; Holmgren, S. K.; Rito, C. J.; Gellman, S. H. *J. Am. Chem. Soc.* **1992**, *114*, 4834. (b) BENZAC02: Feld, R.; Lehmann, M. S.; Muir, K. W.; Speakman, J. C. *Z. Kristallogr.* **1981**, *157*, 215. (c) BAGWUB: Smith, G.; Kennard, C. H. L. *Acta Crystallogr., Sect. B* **1981**, *37*, 1891. (d) BAGXAI: Smith, G.; Kennard, C. H. L. *Acta Crystallogr., Sect. B* **1981**, *37*, 1891.

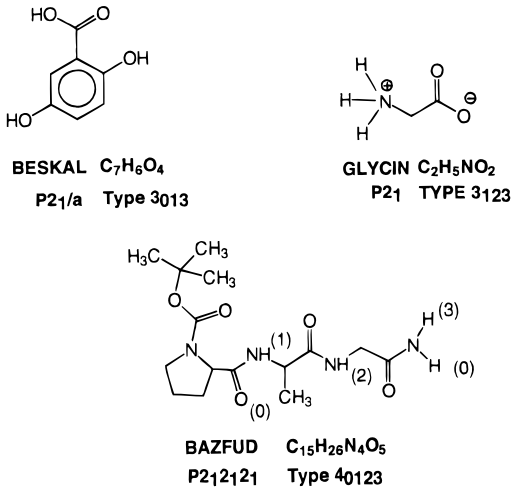


Figure 4. Molecular structures, formulas, space groups, and subtypes belonging to type 3 and type 4 hydrogen-bonded crystals. Not shown are types 3_{012} and 3_{023}. Refcodes are from the Cambridge Structural Database.[27]

Table 2. Energy Components of KAP Stages for H-bond Type 1 (kcal/mol)

KAP stage	Vdw energy	Vdw (H-bond)	Coulomb	total energy
		PAGTIA Type 1_0		
0		intramolecular H-bond		
1	−7.94	0	−2.36	−10.30
2	−15.35	0	−4.58	−19.93
3	−31.93	0	−7.14	−39.07
		BENZAC02 Type 1_1		
0		no H-bond		
1	−4.86	+0.94	−6.71	−10.63
2	−10.76	+0.93	−7.10	−16.93
3	−14.41	+0.92	−7.20	−20.69
		BAGWUB Type 1_2		
0		no H-bond		
1	−8.99	0	−2.49	−11.48
2	−16.30	+1.20	−8.06	−23.16
3	−30.74	+1.20	−8.57	−38.11
		BAGXAI Type 1_3		
0		no H-bond		
1	−12.45	0	−2.04	−14.49
2	−24.26	0	−1.76	−26.02
3	−33.45	+1.25	−6.96	−39.16

have in common that hydrogen bonding occurs in only one stage of KAP. Hydrogen bonding can be considered a vector for that stage and that stage only. The remainder of the packing pattern is determined mainly by the nonbonded potential of eq 1.

1. Type 1_0: PAGTIA. This oligomer has only one intramolecular hydrogen bond in a nine-membered ring as shown visually in Figure 5. Once this intramolecular H-bond is formed, the rest of the crystal packing is controlled by other nonbonded forces. In Table 2, where the breakdown of the energy for each stage is shown, the H-bond contributions to stages 1−3 are all 0. The stage 1−3 packing is thus controlled by the Vdw coupling of the molecules.

2. Type 1_1: BENZAC02. This type is very typical of many simple carboxylic acids.[25] The energy contribution to each stage is shown in Table 2. The H-bond contribution occurs in a stage 1 inversion chain along the b-unit-cell axis with a 5.128 Å repeat as shown in Figure 6 and does not increase any further in stages 2 and 3. H-bonding is thus a vector for stage 1 only. In Figure 6 it can be seen that the chain contains molecules related by inversion on either side of the inversion axis, giving the appearance of a double chain. We also point out here that no matter how strong the H-bond is in stage 1, it will have no effect on the packing energy of stage 2 or 3. We believe this

(25) For a thorough discussion of carboxylic acids and amides, see: (a) Leiserowitz, L. *Acta Crystallogr.* **1976**, *B32*, 775−802. (b) Leiserowitz, L.; Hagler, A. T. *Proc. R. Soc. London, A* **1983**, *388*, 133−175.

Table 3. Energy Components of KAP Stages for H-bond Type 2 (kcal/mol)

KAP stage	Vdw energy	Vdw (H-bond)	Coulomb	total energy
		AMBACO06 Type 2_{01}		
0		intramolecular H-bond		
1	−4.12	+1.84	−7.57	−9.85
2	−8.56	+1.83	−8.08	−14.81
3	−14.09	+1.82	−8.67	−20.94
		FURDCB01 Type 2_{02}		
0		intramolecular H-bond		
1	−9.09	−0.02	−2.32	−11.43
2	−10.11	+2.00	−10.68	−18.79
3	−19.32	+2.00	−11.76	−29.08
		SOJBIC Type 2_{12}		
0		no H-bond		
1	−4.00	+0.32	−6.18	−9.86
2	−13.07	+0.45	−8.72	−21.34
3	−17.86	+0.45	−9.28	−26.69
		METCYT01 Type 2_{13}		
0		no H-bond		
1	−3.40	+0.04	−5.00	−8.36
2	−9.26	+0.03	−5.73	−14.96
3	−14.76	+0.22	−9.16	−23.70

Table 4. Energy Components of KAP Stages for H-bond Types 3 and 4 (kcal/mol)

KAP stage	Vdw energy	Vdw (H-bond)	Coulomb	total energy
		BESKAL Type 3_{013}		
0		intramolecular H-bond		
1	−8.04	+1.08	−6.99	−13.95
2	−16.30	+1.07	−7.99	−23.22
3	−20.04	+1.52	−12.72	−31.24
		GLYCIN Type 3_{123}		
0		no H-bond		
1	−0.61	+0.76	−24.28	−24.13
2	−1.59	+0.87	−36.58	−37.30
3	−6.91	+1.41	−45.40	−50.90
		BAZFUD Type 4_{0123}		
0		intramolecular H-bond		
1	−7.39	+0.43	−8.71	−15.67
2	−13.62	+0.83	−15.29	−28.08
3	−30.96	+1.07	−23.16	−53.05

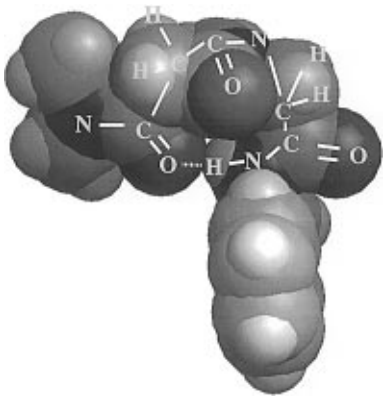


Figure 5. van der Waals surface for stage 0 structure of PAGTIA showing the nine-membered intramolecular hydrogen-bonded ring (type 1_0). There are no other hydrogen bonds in successive stages of the crystal packing.

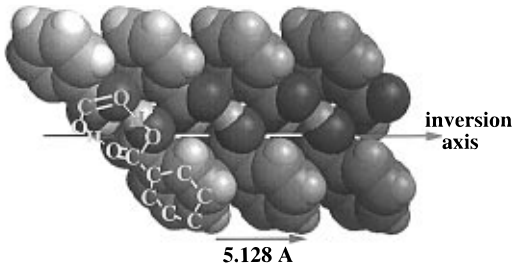


Figure 6. Lowest energy stage 1 packing of benzoic acid (type 1_1). The molecules form a two-molecule-thick inversion chain with hydrogen bond vectors across the inversion axis. There are no additional hydrogen bonds in successive stages of the crystal packing.

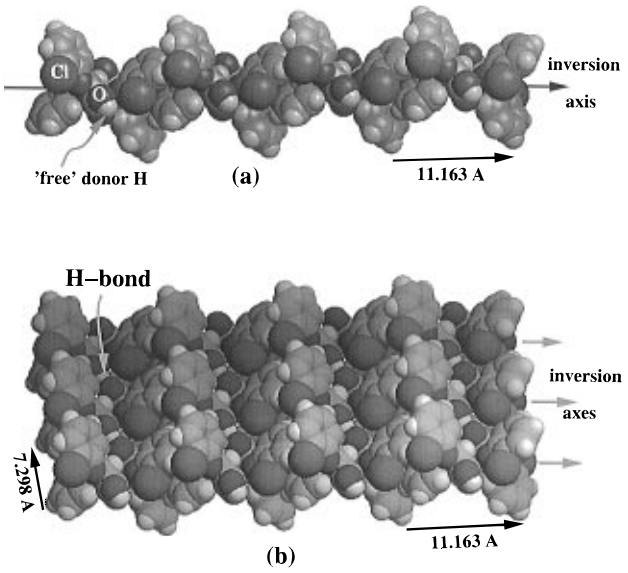


Figure 7. Lowest energy stage 1 and stage 2 packing of BAGWUB (type 1_2). (a) In stage 1, no hydrogen bonds are formed by the inversion chain. These chains couple together in (b) stage 2, where the hydrogen bond vectors first appear. There are no further H-bonds formed in stage 3.

to be a general result. A hydrogen bond, no matter how strong a vector in one stage, can have no effect on any later stages. The packing geometry for the stage 2 and stage 3 interactions will thus not change with an increase in strength of stage 1 H-bonds.

3. Type 1_2: BAGWUB. The H-bond in this structure occurs only in stage 2, as indicated by the energy contributions in Table 2. The H-bond contribution is 0 in stage 1, becomes positive in stage 2, and does not increase in stage 3. Figure 7 shows stages 1 and 2. The H atom from the O donor is "free" in the stage 1 inversion chain, but not in the stage 2 layer. Note the corrugated features of this layer. Again, no matter how strong this stage 2 vector is, it can have no effect on stage 3.

4. Type 1_3: BAGXAI. The H-bond vector occurs only in the last stage of crystal packing, as demonstrated in Table 2. Clearly any change in the strength of this vector will have a dramatic effect on all earlier stages. This type of H-bond would generally be expected to make its appearance in long molecules with a donor atom at one end.

C. Type 2 H-Bond Vectors. This vector type appears in two KAP stages. There are six possibilities, of which we discuss the four examples shown in Figure 3.[26]

1. TYPE 2_{01}: AMBACO06. Anthranilic acid has three donor hydrogens and three acceptor sites. This particular polymorph is an example wherein one donor hydrogen is never used in H-bonding in violation of Etter's rule. There is an

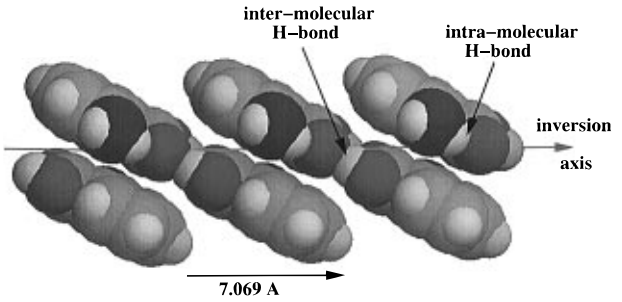


Figure 8. Lowest energy stage 1 inversion chain of AMBACO06 (type 2_{01}). An intramolecular hydrogen bond occurs in stage 0, and intermolecular hydrogen bonds form across the inversion axis in stage 1.

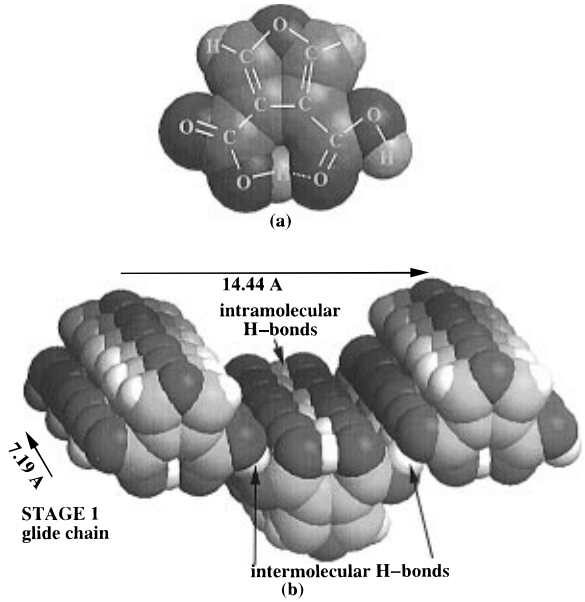


Figure 9. The type 2_{02} H-bond vectors in FURDCB01: (a) stage 0 showing the intramolecular hydrogen-bonded seven-membered ring; (b) stage 1 and stage 2. There are no hydrogen bonds in the lowest energy stage 1 chain. This is a glide chain with a 7.19 Å repeat. These chains couple by screw symmetry in stage 2 with H-bond vectors between the chains producing a corrugated layer with a 14.44 Å repeat.

intramolecular H-bond vector in stage 0 as seen visually in Figure 8, and then another one appears in the inversion chain of stage 1 as indicated by the positive E^{nb}(H-bond) signature in Table 3 and visually in Figure 8. Stages 2 and 3 have no further H-bonds, leaving one H atom on the N donor free. As in BENZAC, the stage 1 chain looks like a two-molecule-thick double chain.

2. Type 2_{02}: FURDCB01. This molecule also has one *intramolecular* H-bond, as shown in Figure 9, but the second H-bond vector does not appear until the layer forms in stage 2. From the energies in Table 3 and visually in Figure 9, the lowest energy stage 1 glide chain with a 7.19 Å c-unit-cell axis repeat vector has no hydrogen bond. The H-bond appears with a large positive Vdw energy in stage 2, and then there is no further increase in stage 3. As seen in Figure 9, the stage 2 packing has a corrugated surface with a long 14.44 Å repeat.

(26) Refcodes are from the Cambridge Structural Database and are as follows. (a) AMBACO06: Hardy, G. E.; Kaska, W. C.; Chandra, B. P.; Zink, J. I. *J. Am. Chem. Soc.* **1981**, *103*, 1074. (b) FURDCB01: Semmingsen, D.; Nordenson, S.; Aasen, A. *Acta Chem. Scand., Ser. A* **1986**, *40*, 559. (c) SOJBIC: Mugnoli, A.; Carnasciali, M. M.; Sancassan, F.; Novi, M.; Petrillo, G. *Acta Crystallogr., Sect. C Cryst. Struct. Commun.* **1991**, *47*, 1916. (d) METCYT01: Rossi, M.; Kistenmacher, T. J. *Acta Crystallogr., Sect. B* **1977**, *33*, 3962.

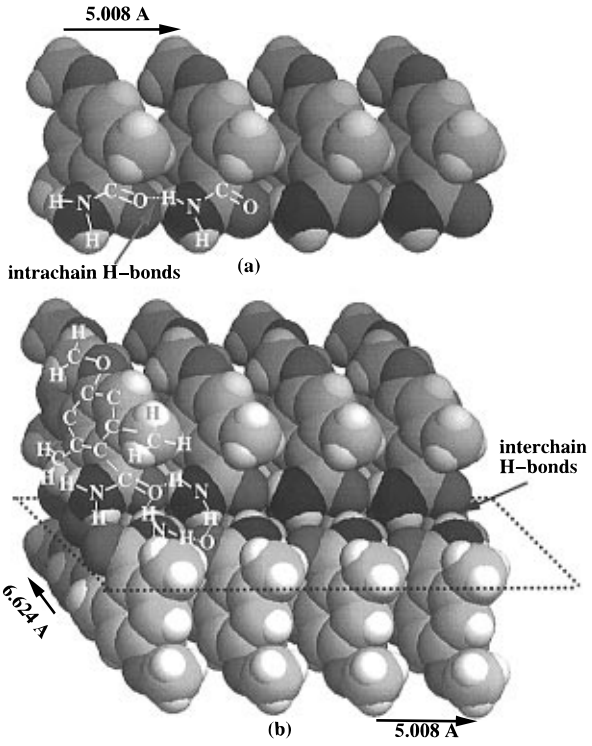


Figure 10. The type 2_{12} H-bond vectors in SOJBIC: (a) a simple translation chain in the lowest energy stage 1 structure with hydrogen bonds formed along the chain; (b) The stage 1 chains couple by inversion symmetry with additional hydrogen bonding in the lowest energy stage 2 layer. Note that the layer is two molecules thick across the layer plane (dotted lines). There are no additional H-bond vectors formed in stage 3.

Again we point out here that, no matter how strong the H-bonding may be in this layer, it will have no effect on the final stage 3 packing since there is no H-bond vector for the final stage.

3. Type 2_{12}: SOJBIC. With two donor hydrogens and potentially three acceptor sites, SOJBIC has two H-bond vectors, one in a stage 1 translation chain with a 5.008 A repeat along the a-unit-cell axis, the other in a stage 2 inversion layer as seen in Figure 10. There is no increase in the H-bond energy in stage 3 (Table 3), so there is no stage 3 H-bond vector. Molecules in the stage 2 layer are related by inversion across the layer plane (dotted line in Figure 10) so that the layer is two molecules thick.

4. Type 2_{13}: METCYT01. Methylcytosine forms an inversion chain with a 8.346 Å repeat in stage 1 using one donor hydrogen as displayed in Figure 11. The repeat vector is not along any unit cell vector but along an a−b diagonal. Since there is no increase in the H-bond energy in stage 2 as seen in Table 3, there is no stage 2 vector. The second H-bond is finally formed in stage 3 as indicated by the positive increase in the H-bond energy.

D. Type 3 H-Bond Vectors. There are four subtypes for type 3 H-bond vectors. We show molecular structural formulas for two examples in Figure 4.[27]

1. Type 3_{013}: BESKAL. BESKAL has three donor hydrogens, and all three are used in hydrogen bonding. In addition to the intramolecular H-bond, Table 4 indicates a positive value

(27) Refcodes are from the Cambridge Structural Database and are as follows. (a) BESKAL: Haisa, M.; Kashino, S.; Hanada, S.-I.; Tanaka, K.; Okazaki, S.; Shibagaki, M. *Acta Crystallogr*., *Sect*. *B* **1982**, *38*, 1480. (b) GLYCIN: Iitaka, Y. *Acta Crystallogr*. **1960**, *13*, 35. (c) BAZFUD: Kojima, T.; Tanaka, I.; Ashida, T. *Acta Crystallogr*., *Sect*. *B* **1982**, *38*, 221.

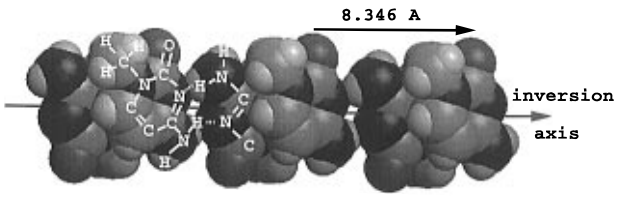


Figure 11. Lowest energy stage 1 structure of METCYT01 with hydrogen bonds formed across the inversion axis. The remaining donor hydrogen atom does not form H-bonds until stage 3.

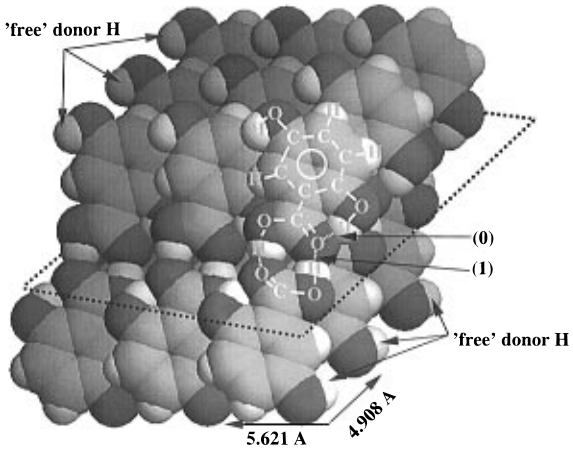


Figure 12. Lowest energy stage 2 layer of BESKAL. There are an intramolecular hydrogen bond in stage 0 (indicated by "(0)"), and intermolecular hydrogen bonds in stage 1 (indicated by "(1)"). Stage 1 is an inversion chain with intermolecular H-bonds formed across the 4.908 Å repeat vector. These chains pack with a repeat of 5.621 Å in stage 2, but no H-bonds are formed in this stage. This is a two-molecule-thick layer (layer plane indicated by dotted lines). The remaining donor hydrogens bond in stage 3.

for the Vdw H-bond energy in stage 1, no increase in this energy in stage 2 and thus no H-bond vector in this stage, and then an increase in this energy in stage 3. This is verified visually in Figure 12. There is an intramolecular H-bond in the stage 0 packing unit containing one molecule. This packing unit then forms a stage 1 inversion chain, two molecules thick with repeat vector 4.908 Å. The second hydrogen bond can be seen in this stage. These chains then pack to form a layer in stage 2, but no additional hydrogen bond is formed in this stage. The remaining donor hydrogen can be seen sitting on top and bottom of this two-molecule-thick layer. These layers then pack together, creating the third H-bond vector in stage 3.

2. Type 3_{123}: GLYCIN. The β form of the amino acid glycine exists as a zwitterion in the solid state. All three donor hydrogens are used in H-bonding in three separate stages as the energy data in Table 4 indicates. In Figure 13, stage 1 is a simple translation chain using the first donor hydrogen with a repeat vector of 5.38 Å. Stage 2 couples these chains by a 2_1 screw symmetry perpendicular to the translation chains producing the layer and making the second H-bond vector with a second repeat vector of 6.268 Å. Finally the third H-bond vector is made by packing these layers to make the full crystal.

E. Type 4 H-Bond Vector. This type should be very common among polypeptides with multiple H-bond possibilities. We present one example shown in Figure 4.[27]

1. Type 4_{0123}: BAZFUD. This polypeptide contains four donor hydrogens, each one used in a single stage of the aufbau. The atoms labeled "(0)" in Figure 4 form an intramolecular H-bond in a 10-membered ring. The donor H atoms labeled

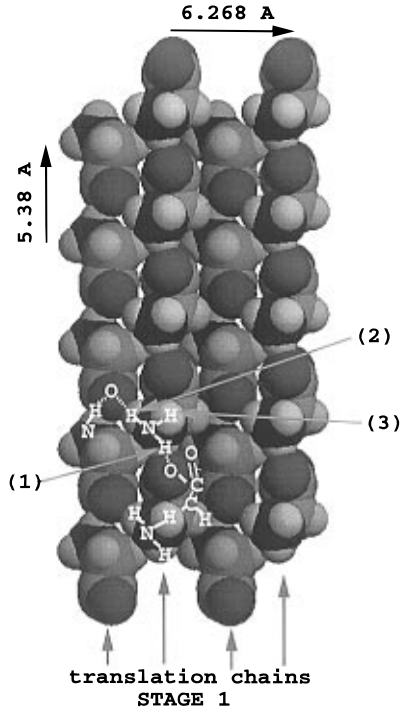


Figure 13. Stage 2 layer structure for GLYCIN (β-glycine) consisting of stage 1 translation chains with a 5.38 Å repeat coupled by screw symmetry to produce a layer with a 6.268 Å repeat. H-bond vectors form in stage 1 (indicated by "(1)"), stage 2 (indicated by "(2)"), and stage 3 (indicated by "(3)" but stage 3 not shown).

1−3 are each used in successive stages 1−3 with a positive increase in the Vdw H-bond energy as indicated in Table 4. Stage 1 is a screw chain with a 10.078 Å repeat (the a-unit-cell axis). Stage 2 contains the stage 1 screw chain repeated by simple translation every 19.246 Å (the b-unit-cell axis). Stage 3 contains the final donor H-bond completing the crystal.

Simulation Studies of H-Bond Types

That the various stages of KAP containing H-bond vectors are in fact local energy minima is easily demonstrated using Monte Carlo simulated annealing. In several recent publications we have shown how to locate the local minima of chains, but we specifically excluded H-bonded ones.[11] Here we show that, with the CFF91 force field, the simulations can also locate the packing geometry of stage 1 H-bonded chains.

For each chain type, we carried out a Monte Carlo cooling procedure on five rigid molecules (viz., molecules with a fixed internal conformation taken from the X-ray structure) initially spaced 15 Å apart in a random orientation. The cooling procedure has been described in great detail elsewhere.[11] The simulation allows the molecules to randomly reorient and translate within the symmetry constraint of the chain type during a slow cooling process from 4000 to 300 K searching for chain structures that have low values of E^{lattice} (eq 3). The cooling process is repeated 700 times, collecting 700 local minima which are sorted for structural uniqueness and rank ordered by energy.[28] Details of the sorting procedure have also been described previously.[11]

Table 5 presents the results for seven H-bonded molecular chains representative of the four important chain symmetries whose molecular structures and Cambridge Structural Database designation are shown in Figure 14.[29] The simulation results show the following features. (a) There are stationary points lying less than 4.25 kcal (column 4) above the apparent global minima which are close to the experimentally observed structures. (b) The root sum square deviation (RSS) values in column 6 of the best local minimum are all small, ranging from 1.21 to 13.03. The RSS represents the quantitative deviation of the molecular orientation and separation distances within the chain.[11] A value of 15.0 for instance for the inversion chains corresponds to angular orientational deviations of less than 5.0° and separation deviations of less than 0.27 Å. (c) The ratios of the electrostatic to the van der Waals terms in column 5 are mostly >1. Thus the electrostatic energy makes a large contribution to the lattice energy of these chains. (d) The predicted separation distances and angles between the H-bonded donor−acceptor atoms shown in columns 8 and 9 are well within the expected range for this type of computation. The largest deviation is 0.1 Å for the distance shown in column 10 and 12.68° for the D−H···A angle shown in column 13.

These results show that H-bonded structures can be included in a natural way into KAP for the construction of chains, layers, and three-dimensional lattices of organic molecules.

Summary and Conclusions

Kitaigorodskii's aufbau principle is a simple but powerful way to understand the crystal structure of arbitrary shaped organic molecules. Inclusion of hydrogen-bonded structures is a natural extension of KAP since most such structures are close packed. We have shown how the aufbau can be viewed as a four-stage process in which hydrogen bonding may occur in one or more stages. Using the CFF91 force field, each stage is a local energy minimum, which we have demonstrated for stage 1 using Monte Carlo simulated annealing techniques. Thus each stage may be considered a separate unique entity. Within that entity, a hydrogen bond may occur as a vector[1e] whose signature is revealed by a positive value for the interaction energy of the hydrogen atom's nonbonded Vdw potential plus repulsive term. We have characterized these vectors as types 1−4, each type representing the number of KAP stages in which an H-bond vector appears. Within each type there are subtypes which give the specific KAP stage combinations for the occurrence of the vectors, and we have presented examples for most of these subtypes. A significant attribute of this typing is that a given vector type cannot have any influence on the packing patterns of any later stages of the aufbau. For crystal engineering this means that the packing features using H-bonds can be designed one stage at a time.

As with all molecular modeling computations, the quality of the results depends on the quality of the force field used. Hydrogen bonding should be particularly sensitive to this since it has a very large electrostatic contribution. Thus we would expect that the details concerning the vector typing might change with change in the electrostatic component of the force field. This does not vitiate our conclusions concerning H-bond vector typing. Provided the same force field is used from one molecule to another, self-consistent results should be attainable. Never-

(28) The Monte Carlo software routines for packing chains, layers, and crystals is called PACK and has been interfaced to the CHEMX/CHEMLIB molecular modeling package. PACK is available upon request. Contact the authors at internet address perlstein@chem.chem.rochester.edu.

(29) Refcodes are from the Cambridge Structural Database. Original references are as follows. (a) SIGBOZ: Zhao, X.; Chang, Y.-L.; Fowler, F. W.; Lauher, J. W. *J. Am. Chem. Soc.* **1990**, *112*, 6627−6634. (b) WARXES: Reference 3f. (c) WARYIX: Reference 3f. (d) DMOXBA01: Bryan, R. F.; White, D. H. *Act Crystallogr.* **1982**, *B38*, 1014−1016. (e) MESCON: Gupta, M. P.; Yadav, S. R. P. *Acta Crystallogr.* **1972**, *B28*, 2682−2686. (f) CITKEV: Smith, G.; O'Reilly, E. J.; Kennard, C. H. L. *Aust. J. Chem.* **1983**, *36*, 2175−2183. (g) DACYEL: Ammon, H. L.; Prasad, S. M. *Acta Crystallogr.* **1985**, *C41*, 921−924.

Table 5. Monte Carlo Cooling Predictions for Hydrogen-Bonded Stage 1 Chains

refc[a]	space group	chain type[b]	ΔE[c]	ratio Coul/Vdw[d]	RSS[e]	rank order[f]	H-bond D···A distances			H-bond D−H···A angles		
							exp[g]	predicted[h]	dev[i]	exp[j]	predicted[k]	dev[l]
SIGBOZ	$P\bar{1}$	t	0.15	1.15	2.45	0	2.915	2.914	−0.001	148.38	149.36	+0.98
							2.917	2.964	+0.047	148.81	149.91	+1.10
WARXES	$P2_1/c$	t	0.81	3.34	1.21	0	2.931	2.912	−0.021	172.45	170.40	−2.05
							2.950	2.923	−0.027	150.96	151.34	+0.38
WARYIX	$C2$	t	0.58	1.33	1.57	0	2.823	2.895	+0.072	144.86	144.38	−0.48
							2.942	2.969	+0.027	132.03	133.69	+1.66
DMOXBA01	$P2_12_12_1$	s	3.68	1.61	8.82	36	2.673	2.777	+0.104	152.01	143.34	−8.67
MESCON	$P2_1/c$	g	0.31	3.14	4.58	0	2.623	2.648	+0.025	165.70	172.15	+6.45
							2.737	2.686	−0.051	161.87	159.14	−2.73
CITKEV	$P\bar{1}$	i	4.24	1.24	8.32	26	2.655	2.724	+0.069	167.39	166.70	−0.69
DACYEL	$P2_1/c$	i	2.41	0.82	13.03	7	2.605	2.661	+0.056	171.44	158.76	−12.68

[a] Reference codes from the Cambridge Structural Database.[29] [b] t = translation, s = screw, g = glide, i = inversion. [c] Energy in kcal/mol of local minimum above the apparent global minimum. [d] Ratio of Coulomb energy to van der Waals energy (eq 2/eq 1). [e] Root sum square deviation of the orientational and translational degrees of freedom of the local minimum from the observed. See ref 11 for details. [f] Number of unique structures whose energy is higher than the apparent global minimum. [g] Data from ref 29 in Å. [h] This work in Å. [i] Predicted minus experimental distances in Å. [j] Data from ref 29 in degrees. [k] This work in degrees. [l] Predicted minus experimental angles in degrees.

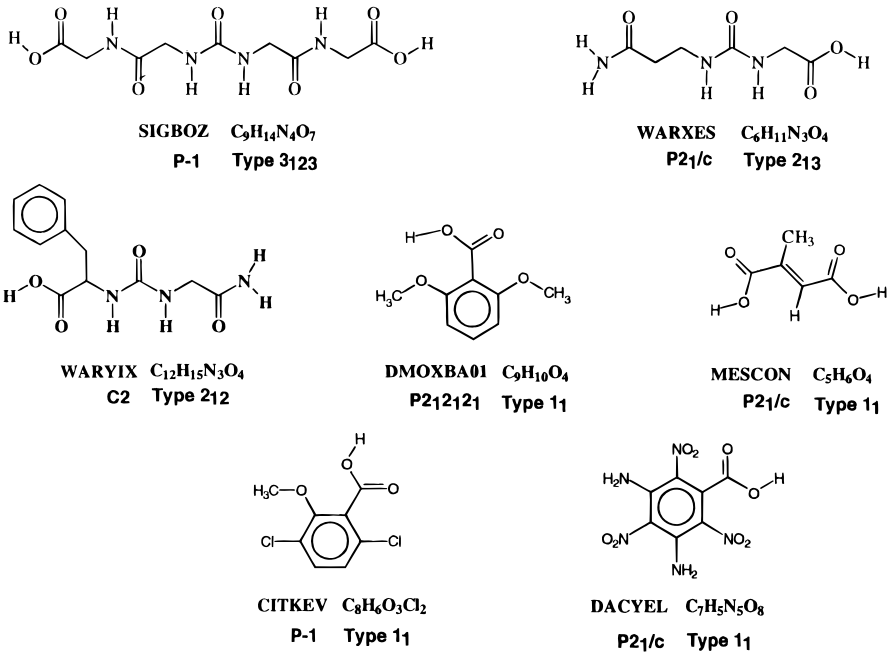


Figure 14. Molecular structures, refcodes,[29] formulas, space groups, and hydrogen bond subtypes of molecules used in Monte Carlo simulated annealing studies. Each molecule has at least a stage 1 substructure with hydrogen bond vectors.

theless it seems worthwhile to explore the electrostatic contribution to the crystal lattice energy where H-bonding is involved. Distributed multipole analysis should be important in this regard and is being explored fruitfully.[30]

There are of course other ways of dissecting the crystal packing of molecules, and other authors have done so.[31] Given the KAP substructures, other kinds of aggregates and clusters can be extracted from them. Most notably, inversion chains often show dimer clusters (see Figures 6, 8, and 11, for example). Other cluster types have also been extracted from glide layers.[32] KAP, however, is the most general of these ways by providing the lowest energy substructures of the infinite lattice.

Vector Typing and Crystal Engineering

In this paper, we have focused our attention on only one vector type, strong and moderately strong hydrogen bonds, and have provided a quantitative signature for it in terms of the CFF91 force field. We have also shown how to find the local energy minima containing this vector type in stage 1 of KAP using Monte Carlo simulation methods.

Other vector types may be present in crystal packing. For instance, various other design elements appear in supramolecular structures which Desiraju calls synthons.[33] Desiraju lists 35 of them, 27 of which interestingly contain the vector type D−H···A in which D and A are donor or acceptor atoms C, N, O. There are other synthons containing other vector types. If quantitative signatures can be found for other vector types which appear in these synthons, then the typing scheme developed in this paper can be applied to each of them. It should then become possible to find the best packing geometry in any stage of KAP by doing

(30) Willock, D. J.; Price, S. L.; Leslie, M.; Catlow, C. R. A. *J. Comput. Chem.* **1995**, *16*, 628−647.

(31) Gavezzotti, A. *J. Am. Chem. Soc.* **1991**, *113*, 4622−4629. For examples of packing in organometallic chemistry, see: (a) Braga, D.; Grepioni, F. *Acc. Chem. Res.* **1994**, *27*, 51−56. (b) Braga, D.; Grepioni, F.; Tedesco, E.; Orpen, A. G. *J. Chem. Soc., Dalton Trans.* **1995**, 1215−1220.

(32) See, for example, the following papers by the Whitten group: (a) Song, X.; Perlstein, J.; Whitten, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 7816−7817. (b) Chen, H.; Law, K.-Y.; Perlstein, J.; Whitten, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 7257−7258. (c) Song, X.; Geiger, C.; Leinhos, U.; Perlstein, J.; Whitten, D. G. *J. Am. Chem. Soc.* **1994**, *116*, 10340−10341.

(33) Desiraju, G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2311−2327.

a Monte Carlo simulation subject to the constraint of the vector typing signature.

Acknowledgment. E.M.N.N. acknowledges financial support from the National Science Foundation through the Educational Outreach Program of the NSF Center for Photoinduced Charge Transfer at the University of Rochester. K.S. participated in the National Science Foundation Research Experience for Undergraduates at the University of Rochester. This research was supported in part by the National Science Foundation Center for Photoinduced Charge Transfer (CHE-9120001).

Supporting Information Available: CFF91 force field parameters for computing *A* and *B* in eq 1 and the *q*'s in eq 2 for atoms used in this work (2 pages). See any current masthead page for ordering and Internet access instructions.

JA9543412