# Molecular Self-Assemblies. 3. Quantitative Predictions for the Packing Geometry of Perylenedicarboximide Translation Aggregates and the Effects of Flexible End Groups. Implications for Monolayers and Three-Dimensional Crystal Structure Predictions

Jerry Perlstein

*Center for Photoinduced Charge Transfer, Department of Chemistry, University of Rochester, Rochester, New York 14627*

The one-dimensional aggregate-packing geometry of 22 perylene dicarboximide pigment molecules consisting of the same rigid chromophore with differing end groups (but taken to be rigid) has been determined using Monte Carlo Cooling techniques to generate local energy minima. The force field used was a nonbonded atom–atom potential taken from MM2 plus a Coulomb electrostatic term with a distance-dependent dielectric constant. The results compare very favorably with experimentally known geometries determined by Klebe et al. In all but four cases there is a local minimum structure which is close to the experimental structure not more than 5 kcal above the apparent global minimum. The number of distinctly different geometries within 5 kcal of the global is never more than nine for each molecular aggregate. The electrostatic energy contribution to the total energy is less than 4%. The packing geometry is thus determined by the nonbonded energy term with changes in the substitution pattern on the ends of the molecules having considerable influence on the overall packing. Inclusion of internal variables (torsion variables) for the end groups in the Monte Carlo simulation has little effect on the local minima determined by the external variables. The results suggest an aufbau principle which we call KAP (Kitaigorodskii's aufbau principle) whose significance for the prediction of monolayer geometries and full three-dimensional crystal geometries of organic molecular self-assemblies with flexible groups is discussed.

## Introduction

Perylenedicarboximides (I) are a member of a class of pigments that consist of two parts, (a) a rigid chromophore having a $\pi$-electronic structure bonded to (b) a generally flexible head and/or tail shown schematically in Figure 1a. This modular design is similar to the description of
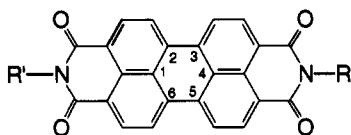


(I)

that for molecules which form bilayer membranes recently reviewed by Kunitake[1] with the notable exception that the perylenedicarboximides are virtually insoluble in most organic solvents. Nevertheless, examination of their crystal structures clearly shows the bilayer nature of the interactions in which the rigid chromophoric parts form stacked aggregate arrays as seen in Figure 1b for R = R′ = *n*-pentyl. The optical and electrical characteristics of chromophoric molecules are dramatically influenced by this stacking behavior.[2] As such these dicarboximides have been actively studied as chromophores for photoinduced electron-transfer processes in the solid state.[3]

One of the interesting features of the stacking behavior of I is the sensitivity of the overlap geometry to the presence of different end groups R. In an extensive X-ray study[4,5] of the effects of 18 different R group substitutions, Klebe et al.[6] concluded that the correlation between the nature of the flexible end group and the type of overlap observed was complicated. Apart from the fact that bulky end groups clearly force the molecules farther apart, the changes in the fine details of the overlap and hence the spectral and electrical characteristics associated with this overlap have been elusive.

We have recently shown that the one-dimensional stacking behavior of most small organic molecules occurs at a local energy minimum of the short-range nonbonded interaction potential not far from the global minimum but structurally very different from it.[7,8] The most important local minima are those in the neighborhood of the global minimum, and these can be found using a Monte Carlo search technique. The four types of stacking behavior most commonly found and methods for constructing them have been described by Scaringe and Perez.[9]

(3) For an extensive review of the literature on the use of perylene dicarboximides as photoconductive materials, see: Borsenberger, P. M.; Weiss, D. S. *Organic Photoreceptors for Xerography*, Marcel Dekker, NY 1993. Law, K.-Y. *Chem. Rev.* 1993, *93*, 449.
(4) Graser, F.; Hadicke, E. *Acta Crystallogr.* 1986, *C42*, 189.
(5) Graser, F.; Hadicke, E. *Acta Crystallogr.* 1986, *C42*, 195.
(6) Klebe, G.; Graser, F.; Hadicke, E.; Berndt, J. *Acta Crystallogr.* 1989, *B45*, 69.
(7) Perlstein, J. *J. Am. Chem. Soc.* 1992, *114*, 1955.
(8) Perlstein, J. *J. Am. Chem. Soc.* 1994, *116*, 455.
(9) Scaringe, R. P.; Perez, S. *J. Phys. Chem.* 1987, *91*, 2394.

● Abstract published in *Advance ACS Abstracts,* February 15, 1994.
(1) Kunitake, T. *Angew. Chem., Int. Ed. Engl.* 1992, *31*, 709.
(2) Kasha, M.; Rawls, H. R.; Ashraf, E.-B. *Pure Appl. Chem.* 1965, *11*, 371.
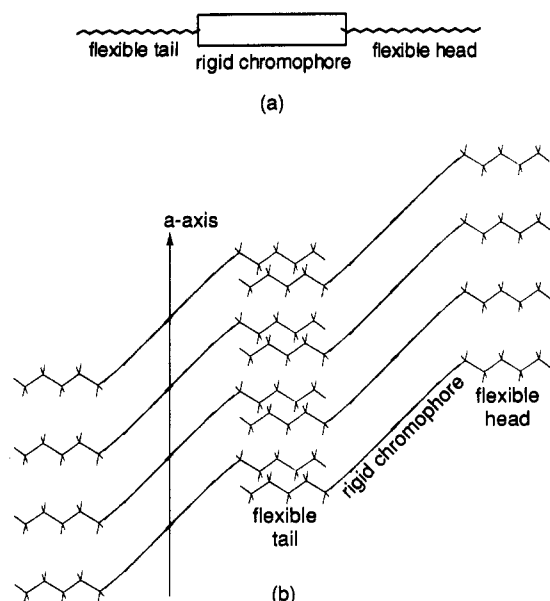
**Figure 1.** (a) Schematic of a rigid chromophore with attached flexible head and tail. This type of model is used extensively in discussions of aqueous bilayer membranes.[1] (b) X-ray structure viewed perpendicular to the short axis of perylene dicarboximide with R = n-pentyl.[4] This view shows the bilayer-like arrangement with the rigid chromophores forming translation aggregates.
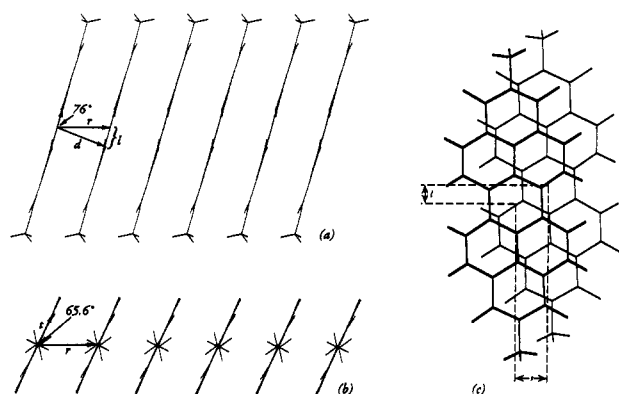


**Figure 2.** (a) X-ray structure of perylene dicarboximide (I) aggregate with R = R′ = methyl. The view is perpendicular to the short molecular axis. The aggregate repeat distance $r = 3.874$ Å, and the interplanar spacing $d = 3.40$ Å. The longitudinal tilt angle between $r$ and the long molecular axis is 76°. (b) Same structure but viewed perpendicular to the long molecular axis. The transverse tilt angle between $r$ and the short molecular axis is 65.6°. (c) Same structure showing only two molecules viewed perpendicular to $d$. As a result of the tilt angles there is a longitudinal slip $l = 0.94$ Å and a transverse slip $t = 1.60$ Å of the molecular mass centers. Compare to Figure 3 with R = propyl.

Graser and Hadicke[4,5] have determined from X-ray structure analysis that all but one of the perylene dicarboximides regardless of substitution form stacked arrays belonging to the simplest type of aggregate structure, namely, the one-dimensional translation aggregate in which every molecule is related to its nearest neighbor in the stack by a simple translation. The effects of the end substituents is to change the tilt angles of the long and the short molecular axis with respect to the translation direction as shown in Figures 2 and 3 for R = methyl and R = propyl.

It is the purpose of this work to demonstrate two things: (a) that the tilt angles and intermolecular spacing are quantitatively predictable given only the rigid valence bond geometry of the molecule as a starting point and (b) that
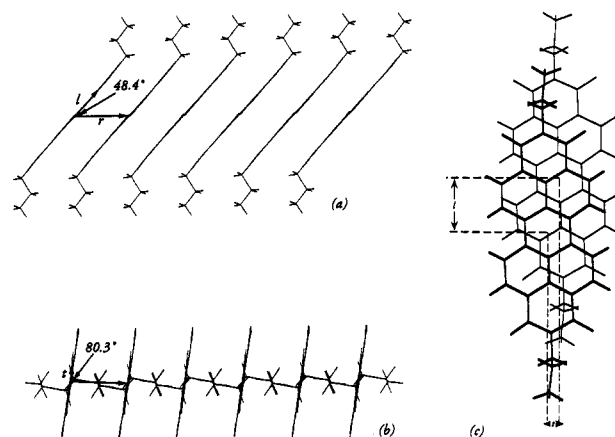


**Figure 3.** X-ray structure of one of the perylene dicarboximide (I) aggregates with R = R′ = propyl showing (a) the longitudinal tilt of 48.4° with a translation repeat $r = 4.689$ Å and an interplanar spacing $d = 3.408$ Å and (b) a transverse tilt of 80.3°. In (c) the molecular tilts cause a longitudinal slip $l = 3.11$ Å and a small transverse slip, $t = 0.79$ Å, as viewed perpendicular to the interplanar spacing. Compare to Figure 2 with R = methyl.

even under conditions in which the valence bond geometry is disturbed by allowing for internal rotations about the single bonds of the flexible end groups, the external tilt angles and molecular spacing are still quantitatively predictable. We believe our technique will contribute to the long outstanding problem of predicting three-dimensional crystal structures without any knowledge of the unit-cell dimensions especially in those situations where the crystal structure determination is difficult to do.[10]

For the rigid molecules we have generated local energy minima for each aggregate using the nonbonded atom–atom potential of MM2 (eq 1)[11]plus an electrostatic term

$$E^{nb} =$$

$$\frac{1}{2}\sum_{i,j}A_{ij}\left[2.90 \times 10^5 \exp\left(\frac{-12.50r_{ij}}{B_{ij}}\right) - 2.25\left(\frac{B_{ij}}{r_{ij}}\right)^6\right] \quad (1)$$

$$E^{el} = \frac{1}{2}\sum_{i,j}\frac{q_iq_j}{\epsilon(r)r_{ij}} \qquad \epsilon(r) = \epsilon_0 r_{ij} \qquad (2)$$

$$E^{tor} = \sum_{\substack{\text{all bond} \\ \text{dihedrals}}} \left\{\frac{V_1}{2}(1 + \cos \omega) + \frac{V_2}{2}(1 - \cos 2\omega) + \right.$$

$$\left. \frac{V_3}{2}(1 + \cos 3\omega)\right\} \quad (3)$$

with a distance-dependent dielectric constant (eq 2).[12] For the flexible molecules we have included the MM2 torsion energy term (eq 3). All parameters (A, B, V) were taken from MACROMODEL version 2.5 which uses the nonbonded MM2 force field equation.[13] The A and B parameters can be found in refs 7 and 8 and the V parameters for eq 3 are listed in Table 5. In eqs 1 and 2 the summation is over all atoms i, of one test molecule interacting with all the other atoms, j, in all the other

(10) Kazmaier, P. M.; McKerrow, A. J.; Buncel, E. *J. Imag. Sci. Technol.* **1992**, *36*, 373.

(11) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

(12) See ref 7 for a description and use of parameters.

(13) Still, W. C.; Mohamadi, F.; Richards, N. B. J.; Guida, W. C.; Lipton, W.; Liskamp, R.; Chang, G.; Hendrickson, T.; DeGunst, F.; Hasel, W. MACROMODEL V2.5, Department of Chemistry, Columbia University, New York, NY 10027.
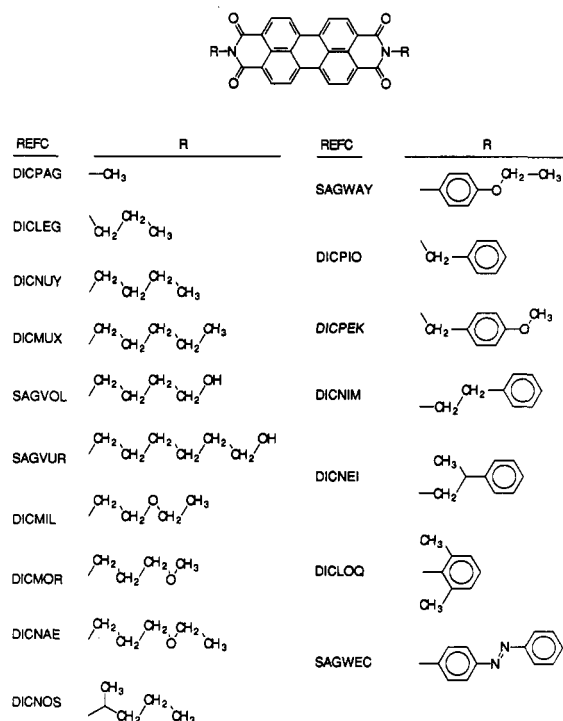
Figure 4. Molecular structures of the perylene dicarboximides. REFC are the reference codes from the Cambridge Structural Database which contains the X-ray atomic coordinates from ref 4–6.

Figure 5. (a) Typical starting geometry used for the Monte Carlo cooling (R = *n*-butyl). Five molecules are randomly oriented with the molecular centroids spaced 15 Å apart. (b) Orientation and spacing for the apparent global minimum after the simulation. Molecules are now 4.47 Å apart with long and short axes tilting with respect to *z*.

molecules in the structure, $q_i$ and $q_j$ are partial atomic Gasteiger charges.[14] $r_{ij}$ is the distance between atoms $i$ and $j$, $\epsilon_0 = 1.0$ and the $^1/_2$ prevents double counting in computing the lattice potential.[15] No hydrogen bond term was included as this particular interaction is virtually nonexistent in these systems (the OH groups in SAGVOL and SAGVUR do not display any propensity for hydrogen bonding).

Starting with the known valence geometry, we generated 350 local energy minima using a Monte Carlo cooling technique written in FORTRAN and described previously.[7] The local minima were then compared with the X-ray results.

## Computational Section

Starting valence geometries for each of the perylene dicarboximide molecules shown in Figure 4 were extracted from the Cambridge Structural Database[16] using CHEM-X software[17] running on a VAX8600. A starting aggregate geometry was constructed using five identical molecules randomly oriented with their centroids lying along the repeat axis (the *z* axis). The centroid of each molecule represented by a dummy atom was computed from the coordinate positions of the six carbon atoms of the central ring (atoms 1–6 in structure I). In addition, the center molecule of the stack had three additional dummy atoms attached to it such that vectors i′, j′, and k′ drawn from the centroid to these three atoms formed an orthogonal molecular coordinate system which rotated with the molecule. Figure 5a shows a typical starting geometry for
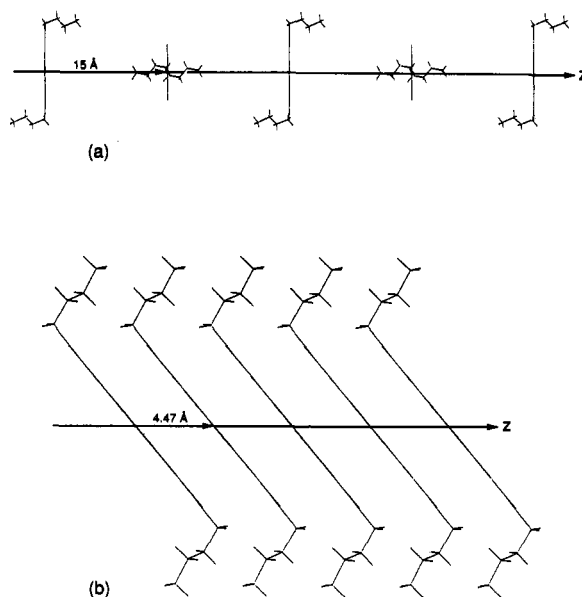
DICNUY, and Figure 5b shows the apparent global minimum generated by the simulation. (The apparent pseudosymmetry shown in Figure 5a is spurious and of no consequence. Any starting geometry can be used since the very first Monte Carlo step will construct an aggregate with translation symmetry). In the starting conformation the molecular coordinate system and the laboratory coordinate system are superimposed. The tilt angles, $a_{i'}$, $a_{j'}, a_{k'}$, that the translation vector makes with the molecular frame vectors i′, j′, k′ are 90°, 90°, 0°, respectively. After the simulation, the tilt angles change. In Figure 5b, for instance, they take on the values 76.53°, 128.55°, and 137.94°. These values, along with the repeat distance $r_{mc}$, contain the complete quantitative information about the packing geometry of the rigid molecules in the aggregate.

**Monte Carlo Cooling.** Monte Carlo cooling was carried out from 4000 to 300 K on an IBM RISC/6000 Model 530 UNIX workstation. (See the flow chart in Figure 7.) For each Monte Carlo step a new aggregate stack was constructed from the old as follows: (a) the center molecule was first randomly rotated about the *x* and *y* axis of the laboratory frame and (b) four identical molecules were then constructed at a random distance of $\pm r_{mc}$, and $\pm 2r_{mc}$ along *z*. The energy of the new aggregate was then computed using eqs 1 and 2, and compared with the old energy using the Metropolis algorithm.[18] The next Monte Carlo step was then done starting with the new state if acceptable or the old state if it was not acceptable. This process was repeated 20 times (10 times for the angles and 10 for the translation) before the temperature was lowered by 10%. At $T = 300$ K the lowest energy state found during the cooling process was saved as a local minimum. The process was then restarted at $T = 4000$ K and repeated

(14) Gasteiger, J.; Marsili, M. *Tetrahedron* 1980, *36*, 3219.

(15) Busing, W. R. *J. Phys. Chem. Solids* 1978, *39*, 691.

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

(17) CHEM-X is a molecular modeling program developed and distributed by Chemical Design Ltd., 7 Westway, Oxford OX2 0JB, U.K.

(18) Metropolis, A. W.; Rosenbluth, M. N.; Teller, A. N.; Teller, E. J. *J. Chem. Phys.* 1953, *21*, 1087. See also ref 7.

(19) Kitaigorodskii, A. I. *Organic Chemical Crystallography*, Consultants Bureau, NY, 1961.

(20) Scaringe, R. P. *Electron Crystallography of Organic Molecules*; Fryer, J. R., Dorset, D. L., Eds.; Kluwer: Dordrecht, 1990.

until 350 local minima were collected. Typical cpu time for a 66-atom molecule was 8 h with an acceptance ratio of 32%.

**X-ray Results.** Geometries of the translation aggregates were extracted from the experimental X-ray data of Graser and Hadicke,[4,5] and Klebe et al.[6] Starting with the asymmetric molecular unit in the unit cell, the structure was replicated with the appropriate space group along the unit cell axes a sufficient number of times so that at least one molecule (the test molecule) was completely surrounded by other molecules. This was conveniently done using CHEM-X software.[17] The interaction potential of this test molecule with each of the other molecules in the structure was then computed using eqs 1 and 2. The lowest energy molecule along with the test molecule then defined a unique aggregate type (translation, screw, glide, or inversion) which for the perylene dicarboximides were translation aggregates except for R = ethyl which forms glide aggregates (but see the discussion concerning SAG-WEC below). Five of the perylenedicarboximides have more than one molecular asymmetric unit in the unit cell. In those cases the lowest energy aggregate for each molecule was used for comparison.

**Comparison of the Monte Carlo with X-ray Data.** Three molecules making up the lowest energy aggregate were extracted from the X-ray structure. Dummy atoms were placed on the center molecule in exactly the same orientation as in the Monte Carlo simulation to define a molecular coordinate frame of reference with unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$. The tilt angles $a_i, a_j, a_k$ of the molecular frame with the translation repeat distance, $r_{\text{X-ray}}$ was computed. This angle data along with the repeat distance was compared to each of the 350 Monte Carlo local minima with tilt angles $a_{i'}, a_{j'}, a_{k'}$ and translation repeat $r_{\text{mc}}$.

We define a root sum square deviation (RSS) of the Monte Carlo data from the X-ray data as follows:

$$RSS = \sqrt{(\alpha_{i'} - \alpha_i)^2 + (\alpha_{j'} - \alpha_j)^2 + (\alpha_{k'} - \alpha_k)^2 + (18r_{\text{mc}} - 18r_{\text{X-ray}})^2}$$

$$(4)$$

where we have weighted the repeat distance by a factor of 18, so that a 0.3-Å difference in repeat distance would have the same weight as a 5.4° difference in tilt angle. These differences would give and RSS of 10.8.

## Results

**Apparent Global Minima.** Table 1 presents the comparison of the apparent global minimum (the lowest energy found by the Monte Carlo cooling] with the X-ray data for each of the translation aggregates. In the table are the reference code for each molecular structure (from Figure 4), the nonbonded energy, coulomb energy, and total energy in kcal/mol for the Monte Carlo structure computed from eqs 1 and 2, and the RSS computed from eq 4.

Of the 22 structures examined, 14 of them have a global minimum with RSS < 20. These are structures whose three tilt angles and translation repeat differ from the observed X-ray structure by no more than 10° and 0.6 Å, respectively. We have pointed out previously[7] that translation aggregates tend to have global minima that are not far from observed geometries and this is borne out here for these perylene dicarboximides. Nevertheless there

**Table 1. Comparison between Global Minimum Structure from the Monte Carlo Cooling and X-ray Data of the Translation Aggregate of Perylene Dicarboximides**

| REFC[a] | nonbonded[b] | coulomb[c] | total[d] | RSS[e] |
|---|---|---|---|---|
| DICPAG | −25.29 | +0.54 | −24.75 | 44.44 |
| DICLEG[f] | −28.36 | − 0.06 | −28.42 | 12.99 |
| | −26.89 | − 0.15 | −27.04 | 28.48 |
| DICNUY | −29.01 | − 0.26 | −29.27 | 5.71 |
| DICMUX | −32.06 | − 0.08 | −32.14 | 16.90 |
| SAGVOL[f] | −31.51 | +0.38 | −31.13 | 13.50 |
| | −31.41 | +0.40 | −31.01 | 14.91 |
| SAGVUR[f] | −35.08 | − 0.13 | −35.21 | 16.74 |
| | −35.24 | +0.06 | −35.18 | 14.97 |
| DICMIL | −28.76 | +0.07 | −28.69 | 6.31 |
| DICMOR | −31.31 | +0.50 | −30.81 | 15.48 |
| DICNAE | −32.28 | − 1.04 | −33.32 | 20.17 |
| DICNOS | −29.46 | − 0.38 | −29.84 | 61.74 |
| SAGWAY[f] | −26.56 | − 1.24 | −27.80 | 81.12 |
| | −26.50 | − 1.41 | −27.91 | 9.05 |
| DICPIO | −30.63 | − 0.31 | −30.94 | 6.07 |
| DICPEK | −37.52 | +0.48 | −37.04 | 7.64 |
| DICNIM | −34.89 | 0.00 | −34.89 | 13.33 |
| DICNEI | −31.21 | − 0.61 | −31.82 | 18.58 |
| DICLOQ[f] | −14.72 | +0.80 | −13.92 | 51.41 |
| | −15.52 | +0.60 | −14.92 | 54.38 |
| SAGWEC | −33.94 | − 0.27 | −34.21 | 82.92 |

[a] REFC are the reference codes from the Cambridge Structural Database.[16] See Figure 4 for structures. [b] Nonbonded energy computed from eq 1. [c] Coulomb electrostatic energy computed from eq 2. [d] Total energy, nonbonded + electrostatic. [e] Root-sum-square deviation of the global minimum from the X-ray structure computed with eq 4. [f] Two distinct translation aggregates associated with two asymmetric molecular units in the unit cell. All energies are in kcal/mol.

are other local minima which are even closer to the observed than the global.

**Local Minima.** Not more than 5 kcal above the apparent global minimum lie local minima which are very close to the observed X-ray structures. These are shown in Table 2. Of the 22 structures examined by the Monte Carlo cooling technique, 19 of them have a local minima whose tilt angles and translation repeat differ from the X-ray data by less than 2.5° and 0.15 Å, respectively, corresponding to RSS values <5. These structures on average are no more than 2 kcal from the global minimum with the worst case being 5.11 kcal for SAGWAY. (But see below concerning SAGWEC.) In all cases the coulomb term (eq 2) is quite small, representing only 4% of the energy in the worst case.

Three of the structures have larger deviations (DICPAG, RSS = 14.42, DICLOQ, RSS = 23.38 and 12.15), and one structure, SAGWEC, has a local minimum which is exceptionally far above the global. We have looked at the observed X-ray structure for each of these in more detail and have found the following:

DICPAG: The observed X-ray structure for this aggregate is not quite at a local minimum, differing from it by a 10° rotation. This is reflected in the somewhat larger RSS of 14.42 for this structure.

DICLOQ: There are two asymmetric molecules forming two distinct translation aggregates. The one with RSS = 23.38 is far from a local minimum both for rotation and translation along the repeat direction. The aggregate with RSS = 12.15 is similar to DICPAG, in deviating somewhat from the bottom of the energy well by a 5° rotation and 0.1 Å in translation.

SAGWEC: This is a most unusual structure. Compared to the global minimum structure in Table 1, the best local minimum is more than 10 kcal higher in energy. This seemed rather large to us compared to the other translation

**Table 2. Comparison between Best Local Minimum Structure from the Monte Carlo Cooling and X-ray Data of the Translation Aggregate of Perylene Dicarboximides**

| REFC[a] | nonbonded[b] | coulomb[c] | total[d] | $\Delta E$[e] | RSS[f] |
|---|---|---|---|---|---|
| DICPAG | −23.95 | +0.42 | −23.53 | 1.22 | 14.42 |
| DICLEG[g] | −26.82 | −0.18 | −27.00 | 1.42 | 2.46 |
| | −24.61 | −0.30 | −24.91 | 2.13 | 4.98 |
| DICNUY | −27.44 | −0.25 | −27.69 | 1.58 | 2.29 |
| DICMUX | −29.92 | −0.23 | −30.15 | 1.99 | 4.40 |
| SAGVOL[g] | −29.13 | +0.18 | −28.95 | 2.18 | 4.36 |
| | −29.47 | +0.06 | −29.41 | 1.60 | 3.18 |
| SAGVUR[g] | −32.13 | −0.36 | −32.49 | 2.72 | 3.98 |
| | −33.62 | −0.25 | −33.86 | 1.32 | 6.36 |
| DICMIL | −28.66 | +0.06 | −28.60 | 0.09 | 5.04 |
| DICMOR | −28.18 | +0.12 | −28.06 | 2.75 | 2.55 |
| DICNAE | −29.44 | −0.86 | −30.30 | 3.02 | 4.46 |
| DICNOS | −26.74 | −0.06 | −26.80 | 3.04 | 3.81 |
| SAGWAY[g] | −21.54 | −1.15 | −22.69 | 5.11 | 1.13 |
| | −24.63 | −1.08 | −25.71 | 2.20 | 3.05 |
| DICPIO | −29.93 | −0.26 | −30.19 | 0.75 | 2.38 |
| DICPEK | −36.86 | +0.33 | −36.53 | 0.51 | 2.78 |
| DICNIM | −33.09 | −0.18 | −33.27 | 1.62 | 2.82 |
| DICNEI | −27.42 | −0.66 | −28.08 | 3.74 | 4.33 |
| DICLOQ[g] | −11.00 | +0.16 | −10.84 | 3.08 | 23.38 |
| | −10.30 | −0.08 | −10.38 | 4.54 | 12.15 |
| SAGWEC | −23.11 | −0.82 | −23.93 | 10.28 | 1.48 |

[a] REFC are the reference codes from the Cambridge Structural Database.[16] See Figure 4 for structures. [b] Nonbonded energy computed from eq 1. [c] Coulomb electrostatic energy computed from eq 2. [d] Total energy, nonbonded + electrostatic. [e] Energy difference between local minimum and global minimum from Table 1. [f] Root-sum-square deviation of the global minimum from the X-ray structure computed with eq 4. [g] Two distinct translation aggregates associated with two asymmetric molecular units in the unit cell. All energies are in kcal/mol.

aggregates we have looked at. The translation aggregate is definitely at the bottom of a local energy well, but this turns out not to be the lowest energy aggregate for this molecule. We reexamined the X-ray data more closely and found that there is an inversion aggregate with an energy of −32.12 kcal and a repeat distance of 12.761 Å along the *a*, *b* face diagonal of the unit cell. This aggregate is some 8.3 kcal lower in energy than the translation aggregate. Because the molecule sits on an inversion point in the unit cell, this inversion aggregate was initially missed since all the molecules are related by translation symmetry.

**Alternative Comparison Method.** Klebe et al.[6] analyzed the stacking behavior for these 22 perylenedicarboximide aggregates from a different perspective. They presented the structures in terms of the overlap of the planes of two neighboring molecules (viz., viewed along the interplanar separation *d*, rather than along the translation repeat). In Figures 2 and 3 for example if one views the structure of one molecule perpendicular to the central ring, the centroid of a neighboring molecule appears offset both longitudinally a distance *l*, and transversely a distance *t*, from the first molecule. If we setup a molecular coordinate system with unit vectors **i′**, **j′**, and **k′** such that **j′** is computed from the vector defined from the centroid to a dicarboximide nitrogen atom, **k′** perpendicular to the six-center carbon atoms and **i′** = **j′** × **k′** as shown in Figure 6 then Klebe et al. values for *l*, *t*, and *d* are related to the direction cosines **i′·r**, **j′·r**, and **k′·r** by the simple relations

$$l = \mathbf{j'} \cdot \mathbf{r} \tag{5}$$

$$t = \mathbf{i'} \cdot \mathbf{r} \tag{6}$$

$$d = \mathbf{k'} \cdot \mathbf{r} \tag{7}$$

where **r** is the translation repeat vector.



$$\mathbf{j'} = \frac{\mathbf{y}}{|\mathbf{y}|}$$

$$\mathbf{k'} = \frac{\mathbf{z}}{|\mathbf{z}|}$$

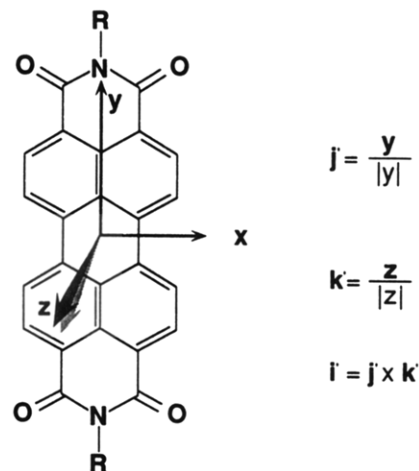$$\mathbf{i'} = \mathbf{j'} \times \mathbf{k'}$$

**Figure 6.** Definition of molecular coordinate system with centroid defined by the 6 central carbon atoms. **k′** is normal to these six atoms, *j′* lies along the y axis defined from the centroid to the imide nitrogen, and **i′** is the vector cross product **j′** × **k′**.
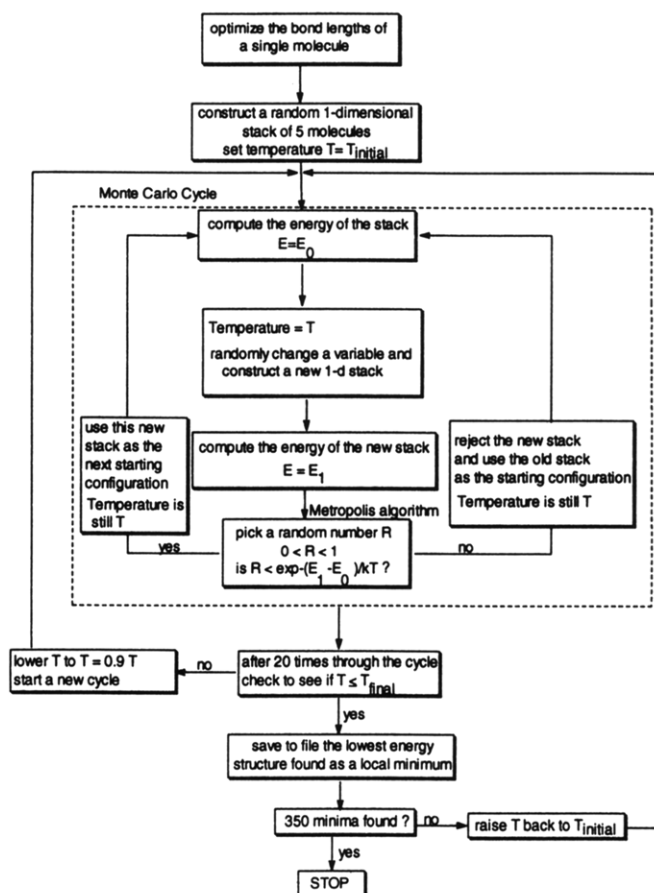


**Figure 7.** Flow chart for the Monte Carlo cooling procedure to generate 350 local energy minima of one-dimensional stacks. $T_{initial}$ = 4000 K, $T_{final}$ = 300 K requiring approxiamtely 354 000 Monte Carlo steps.

Using the local minima from Table 2 and leaving out DICPAG and DICLOQ, we have computed *l*, *t*, and *d* and compared them with Klebe et al.[6] results in Table 3. The differences shown in columns 8–10 of the table are small, the worst case being 0.395 Å difference for the interplanar spacing of SAGWAY.

It is interesting to note that for all the structures studied, the major intermolecular interaction is between neighboring perylene rings. This is because the flexible end groups contain less atoms compared to the rigid perylene

324 *Chem. Mater., Vol. 6, No. 3, 1994*

*Perlstein*

**Table 3. Comparison between the Best Local Minima and the Observed X-ray Structure of the Interplanar Spacing, d, Longitudinal and Transverse Slip Distances, l and t**

| REFC[a] | Monte Carlo[b] | | | X-ray[c] | | | differences[d] | | |
|---|---|---|---|---|---|---|---|---|---|
| | d | l | t | d | l | t | | | |
| DICLEG[e] | 3.309 | 1.79 | 2.35 | 3.411 | 2.05 | 2.48 | -0.102 | -0.26 | -0.13 |
| | 3.288 | 3.10 | 0.71 | 3.408 | 3.11 | 0.79 | -0.120 | -0.01 | -0.08 |
| DICNUY | 3.379 | 3.13 | 1.19 | 3.402 | 3.10 | 1.11 | -0.023 | 0.03 | 0.08 |
| DICMUX | 3.270 | 2.87 | 1.21 | 3.343 | 3.11 | 1.31 | -0.073 | -0.24 | -0.10 |
| SAGVOL[e] | 3.286 | 3.17 | 0.66 | 3.355 | 3.22 | 0.76 | -0.069 | -0.05 | -0.10 |
| | 3.309 | 3.02 | 0.80 | 3.510 | 3.06 | 0.70 | -0.200 | -0.04 | 0.10 |
| SAGVUR[e] | 3.349 | 3.17 | 0.57 | 3.415 | 3.13 | 0.85 | -0.066 | 0.04 | -0.28 |
| | 3.220 | 2.83 | 0.87 | 3.395 | 3.19 | 1.10 | -0.175 | -0.36 | -0.23 |
| DICMIL | 3.374 | 2.92 | 1.10 | 3.355 | 3.23 | 1.28 | 0.019 | -0.31 | -0.18 |
| DICMOR | 3.482 | 3.10 | 0.84 | 3.455 | 3.15 | 0.68 | 0.027 | -0.05 | 0.16 |
| DICNAE | 3.257 | 3.16 | 0.98 | 3.402 | 3.06 | 1.26 | -0.145 | 0.10 | -0.28 |
| DICNOS | 3.337 | 5.24 | 1.26 | 3.422 | 5.45 | 0.98 | -0.045 | -0.21 | 0.28 |
| SAGWAY[e] | 3.475 | 0.44 | 3.50 | 3.870 | 0.55 | 3.50 | -0.395 | -0.11 | 0.00 |
| | 3.452 | 1.43 | 3.04 | 3.550 | 1.56 | 3.09 | -0.098 | -0.13 | -0.05 |
| DICPIO | 3.396 | 2.98 | 1.04 | 3.425 | 3.08 | 1.10 | -0.029 | -0.10 | -0.06 |
| DICPEK | 3.349 | 2.57 | 0.31 | 3.459 | 2.67 | 0.36 | -0.110 | -0.10 | -0.05 |
| DICNIM | 3.330 | 3.19 | 0.40 | 3.476 | 3.20 | 0.39 | -0.146 | -0.01 | 0.01 |
| DICNEI | 3.317 | 1.36 | 3.25 | 3.426 | 1.23 | 3.40 | -0.109 | 0.13 | -0.15 |
| SAGWEC | 3.416 | 0.79 | 3.84 | 3.480 | 0.81 | 3.96 | -0.064 | -0.02 | -0.12 |

[a] REFC are the reference codes from the Cambridge Structural Database.[16] See Figure 4 for structures. [b] Data in angstroms computed from the Monte Carlo local minima (see eqs 5–7 and Figures 2 and 6 for definitions of d, l, and t). [c] Data in angstroms computed from the lowest energy translation aggregate of the X-ray structure.[4–6] [d] Computed differences for d, l, and t between Monte Carlo minimum and X-ray data. [e] Two distinct translation aggregates associated with two asymmetric molecular units in the unit cell.

**Table 4. Rank Ordering of the Local Minima. Number of Unique Structures within 5 kcal of the Global Minimum**

| REFC[a] | unique structures[b] | position of local min[c] |
|---|---|---|
| DICPAG | 14 | 7 |
| DICLEG[d] | 6 | 3 |
| | 9 | 3 |
| DICNUY | 5 | 0 |
| DICMUX | 6 | 2 |
| SAGVOL[d] | 5 | 3 |
| | 6 | 1 |
| SAGVUR[d] | 6 | 1 |
| DICMIL | 7 | 0 |
| DICMOR | 8 | 1 |
| DICNAE | 4 | 3 |
| DICNOS | 9 | 8 |
| SAGWAY[d] | 3 | 2 |
| | 6 | 0 |
| DICPIO | 4 | 0 |
| DICPEK | 4 | 0 |
| DICNIM | 6 | 1 |
| DICNEI | 5 | 2 |
| DICLOQ[d] | 25 | not a local minimum |
| | 24 | 23 |
| SAGWEC | 5 | >5 kcal above global[e] |

[a] REFC are the reference codes from the Cambridge Structural Database.[16] See Figure 4 for structures. [b] Number of unique structures with RSS < 10.0; see text for discussion. [c] Rank order of the best Monte Carlo local minimum beginning with 0. [d] Two distinct translation aggregates associated with two asymmetric molecular units in the unit cell. [e] Lowest energy stacking is not a translation aggregate. See discussion in text.

chromophore. (There are 40 atoms in the chromophore and a maximum 23 atoms in the end group of SAGWEC). Since the interaction potential goes as $N^2$ where $N$ is the number of atoms, this observation should not be surprising. Presumably with considerably longer end groups (such as a $C_{18}H_{36}$ chain) a considerable increase in the tilt angle might be expected which would break up the rigid chromophore interaction. This computation, however, is well beyond both the atom limit and cpu time that we can handle at the moment.

**How Many Unique Local Minima Are There?** We have collected 350 local minima for each structure, but not all of these minima are unique. To determine the number of different structures, we will assume that two structures are the same if the RSS between them is less than 10.0. This means that their tilt angles and translation repeat on average do not differ by more than 5° and 0.28 Å, respectively. To find these unique structures we did the following:

(1) The structures were ordered by energy with the global minimum at the top of the list.

(2) Starting with the global minimum as the first unique structure, the remaining structures were compared by computing the RSS with respect to the global. All of those that had RSS < 10.0 were considered similar to the global. These are the zeroth-order structures. They were removed from the list.

(3) The structure now at the top of the list became the next unique structure. The remaining structures were compared to it. Those with RSS < 10.0 became the 1st order structures and were removed from the list.

(4) Step 3 was repeated until the list was exhausted, finding the 2nd, 3rd, etc. order structures.

We carried out this procedure for all the perylene dicarboximides. The results are shown in Table 4. Within 5 kcal of the global minimum the number of unique structures varies from a low of 3 for SAGWAY to a high

of 25 for DICLOQ with most molecules having no more than 4–6. The position of the best local minimum is no worse than order 3 in the majority of cases.

**Inclusion of Torsion Terms.** The presence of internal variables complicates the molecular packing problem enormously. This can be understood simply by the fact that the van der Waals surface topology of the molecule is different for each of the possible molecular conformations. Consequently the way the molecules pack together will be different for each conformation. It is like trying to put a jigsaw puzzle together while the shape of the pieces keeps changing.

Take for example DICNUY with eight single bonds outside the rigid chromophore. Even if there were only three possible conformations for each bond, this would amount to $3^8$ or approximately $10^4$ different starting conformations. There are approximately 75 000 translation aggregates that are possible for each conformation. This results in a conformation space of some $7.5 \times 10^8$ aggregate structures. We cannot hope to find all the important local minima in a reasonable time frame much less the single one that happens to occur in the X-ray crystal structure.

A more reasonable question to ask is the following: How critical is the shape of the molecule for determining the external variable aggregate geometry (viz. the tilt angles and repeat distance)? Put another way, Are there a sufficiently large number of conformations which give the same aggregate tilt angles and repeat length and can we still find the external variable geometry as a local minimum even in the presence of a large number of internal degrees of freedom?

We find in fact that the external variable geometry of the experimentally observed aggregates (and many more that we have studied)[21] still occur as local minima in the simulations even when we include rotations about the

(21) Perlstein, J. *Nature*, submitted.

**Table 5. Torsional Parameters $V_1$, $V_2$, and $V_3$ for DICNIM and DICNUY Used in Eq 3[a]**

| bond and atom types[b] | $V_1$ | $V_2$ | $V_3$ |
|---|---|---|---|
| C2—N2—C3—C3 | 0 | 0 | 0 |
| C2—N2—C3—H | 0 | 0 | -0.2400 |
| N2—C3—C3—C2 | 0 | 0 | 0.1800 |
| N2—C3—C3—H | 0 | 0 | 0.400 |
| N2—C3—C3—C3 | 0 | 0 | 0.400 |
| H—C3—C3—C2 | 0 | 0 | 0.500 |
| H—C3—C3—H | 0 | 0 | 0.2370 |
| H—C3—C3—C3 | 0 | 0 | 0.2670 |
| C3—C3—C3—C3 | 0.2000 | 0.2700 | 0.6000 |
| C3—C3—C2=C2 | -0.4400 | 0.2400 | 0.0600 |
| H—C3—C2=C2 | 0 | 0 | -0.2400 |

[a] Parameters are from MACROMODEL[13] in kcal/mol. [b] Atom types are C2 = $C_{sp_2}$, C3 = $C_{sp_3}$, and N2 = $N_{sp_2}$.

**Table 6. Comparison between the Minima for DICNIM and DICNUY When Internal Torsion Angles Are Included as Variables**

| REFC[a] | RSS (external)[b] | RMSD (internal)[c] | rank order[d] |
|---|---|---|---|
| DICNIM |  |  |  |
| initial | 59.47 | 73.55 |  |
| global | 12.66 | 9.23 | 0 |
| local | 2.13 | 12.94 | 3 |
| DICNUY |  |  |  |
| initial | 61.75 | 104.42 |  |
| global | 10.24 | 58.68 | 0 |

[a] Reference codes from the Cambridge Structural Database (see Figure 4). [b] RSS computed from equation 4. [c] RMSD computed from eq 9. [d] Number of structures with lower energy.

single bonds as variables in the Monte Carlo cooling. We give only brief details here for two of the molecules DICNIM and DICNUY and will present complete details for these as well as for many other types in a later paper.

DICNIM has six internal rotational degrees of freedom with two phenyl rings at opposite ends of the dicarboximide chromophore whereas DICNUY has eight internal rotational degrees of freedom from two saturated alkane chains. The initial conformational starting geometry was completely random and far from the observed.

Each Monte Carlo step consisted of a random search on the external variables as described previously plus a random search on the internal variables taken one at a time. The total energy $E_t$ after each step was computed from eq 8, where $E^{nb}$ and $E^{el}$ represent the energy con-

$$E_t = E^{nb} + E^{el} + (E^{int} - E^\circ) \qquad (8)$$

tribution given by changes in the external variables and are given by eqs 2 and 3 and $E^{int}$ represents the internal energy contribution referenced to the energy of the starting conformation given by $E^\circ$. The internal energy consists of three terms: (a) the nonbonded energy of all atoms that are at least 1–4 and given by an expression equivalent to eq 1 with atoms $i$ and $j$ being in the same molecule and $i \neq j$; (b) an electrostatic term equivalent to eq 2; (c) a torsional energy term given by eq 3. The $V$ parameters in eq 3 are from MACROMODEL and are listed in Table 5. In eq 3, the summation is over all dihedral angles for each rotating bond.

The results of the Monte Carlo cooling simulations for these two molecules are shown in Table 6 where the RSS deviations for the starting geometry, the global and best local minimum from the observed X-ray structure are shown along with the root-mean-square deviations of the backbone dihedral angles (RMSD). RMSD was computed from eq 9, where $\omega_i$ are the $N$ backbone dihedral angles

(8 for DICNUY, 6 for DICNIM) and $\omega_0$ are the observed X-ray structural values.

$$RMSD = \sqrt{\sum_i (\omega_i - \omega_0)^2 / N} \qquad (9)$$

Both starting geometries are far from the observed as indicated by the large RSS and RMSD values for the initial structures. For DICNIM, the Monte Carlo simulation finds both a global and a local minima which are close to the observed structure including the internal torsion angles. Thus the large flat rigid phenyl ring at the ends of the molecule considerably reduce the number of torsional conformations that are allowed in the aggregate geometry. DICNUY on the other hand has a global minimum (which is also the best local minimum with rank order 0) for the external variables which is close to the observed (RSS = 10.24) but the internal torsional geometry is still very far from the observed (RMSD = 58.68). There are many internal conformations for this molecule which are allowable with the same external geometry. In both cases the Monte Carlo cooling finds the observed structure as a local minimum even when torsion angles are included as variables.

## Summary and Conclusions

The packing of 22 rigid perylene dicarboximide molecules into one-dimensional translation aggregates has been studied in detail using Monte Carlo cooling techniques to find the apparent global minimum and nearby local minima for each structure using only a nonbonded atom–atom potential taken from the MM2 force field and a Coulomb electrostatic term with a distance dependent dielectric constant to compute the energy. In all but four cases a local energy minimum was found lying no more than 5 kcal above the apparent global minimum which compared quite favorably with the extensive experimental X-ray data of Graser and Hadicke[4,5] and Klebe et al.[6] In 14 cases the global minimum itself was a reasonable representation of the structure. Of the 350 local minima collected for each molecule, there were only nine or less unique structures within 5 kcal of the global minimum for the majority of them. The four molecules that were not satisfactorily searched either (a) had structures that did not lie at a local minimum or (b) in one case (SAGWEC) the translation aggregate was not the lowest energy structure.

Inclusion of torsions angles in the simulations which gives the molecules flexibility yields essentially the same results for the packing of the aggregate structures.

The conclusions to be drawn from this study are as follows: (a) for rigid molecules of known valence geometry, the most important local minima of the one-dimensional aggregate structures including the fine details of the overlap geometry are readily found by a Monte Carlo cooling technique; (b) these structures are close packed and thus controlled almost exclusively by the nonbonded nonelectrostatic atom–atom potential; (c) inclusion of flexible side groups in the simulation still reproduces the important local packing minima.

**Significance of the Results.** The significance of these results, especially point (c), relates to the ultimate predictability of organic monolayer geometries and full three-dimensional organic molecular crystal structures.

There is an aufbau principle here which we dub KAP (Kitaigorodskii's aufbau principle).[19] The utility of KAP is to significantly reduce the phase space both for the internal and external variables that has to be searched. It suggests that the local minima structures for monolayers can be found by first finding the one-dimensional aggregate local minima including torsion angles as variables and then packing these structures together to form layers. The results should be a collection of local minima layers the lowest energy ones of which will be those that occur in Langmuir–Blodgett films, or as epitaxial films on metal surfaces, or as the substructures in the three-dimensional crystal structure of the molecule, or in any other environments for which layer geometries need to be molecularly engineered. KAP can then be used a third time to pack the local energy minima of the layers together to form multilayer structures including the full three-dimensional crystal structure.

We are presently using KAP to predict monolayer geometries[20] of molecules including internal torsion angles as variables and will report on those results shortly.[21]