Computations and ^{1}H NMR Spectroscopy of the Imide Region Can Distinguish Isomers of Hydrogen-Bonded Aggregates

**Donovan N. Chin, Eric E. Simanek, Xinhua Li,
Mohammed I. M. Wazeer, and George M. Whitesides**

Department of Chemistry and Chemical Biology, Harvard University,
12 Oxford Street, Cambridge, Massachusetts 02138

The Journal of

Organic

Chemistry®

Computations and 1H NMR Spectroscopy of the Imide Region Can Distinguish Isomers of Hydrogen-Bonded Aggregates

Donovan N. Chin, Eric E. Simanek, Xinhua Li, Mohammed I. M. Wazeer, and George M. Whitesides*

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138
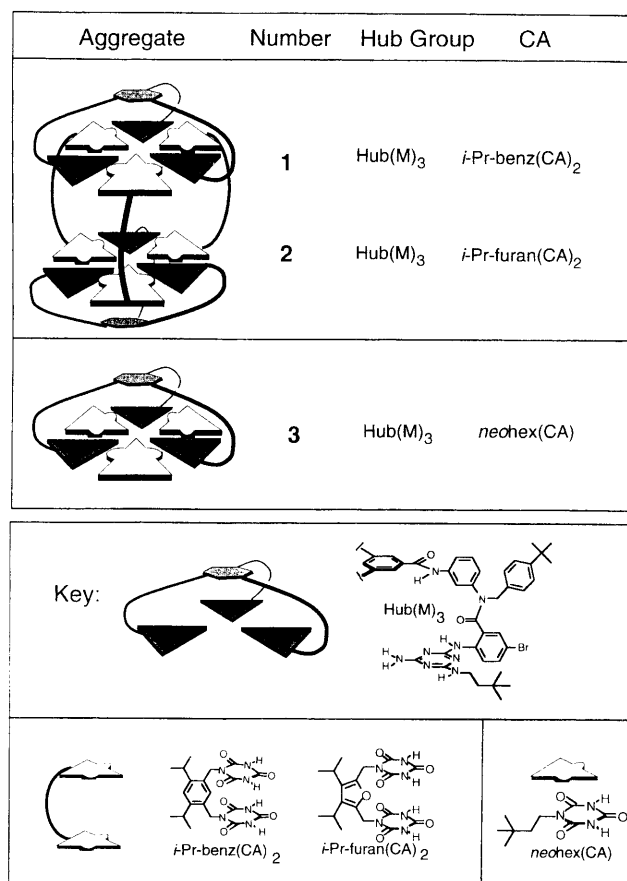
Introduction

This work reports the use of computations and 1H NMR spectroscopy for the structural assignment of isomers of hydrogen-bonded aggregates in solution. Analysis of the hydrogen-bonding imide region of the 1H NMR spectrum (between δ 16 and 13 ppm) shows the number and symmetries of isomers of aggregates based on the isocyanuric acid·melamine rosette.[1] This region is much simpler than the aromatic or aliphatic regions of the spectrum since it is clear of other lines. We have reported rules for analysis of this region previously using a model system: hub(M)$_3$·3neohexylCA, **3** (Scheme 1).[1] As aggregates become increasingly large—further complicating the 1H NMR spectrum—understanding the generality of these rules becomes increasingly important. Analysis of aggregates **1** and **2** suggests that both aggregates show the expected structural behavior and that steric repulsion between groups along the periphery—identified using computation—leads to fewer numbers of isomers than expected: for **1**, two instead of nine. The computational surrogate for relative stability, DP (DP = deviation from planarity),[2–4] is consistent with this assignment.

Hub(M)$_3$·3CA as a Model System. Hub(M)$_3$·3CA (**3**) and aggregates **1** and **2** are based on a cyclic motif of alternating molecules of isocyanuric acid and melamine, which we refer to as a rosette. Each rosette is held together by 18 hydrogen bonds. Hub(M)$_3$·3CA is the simplest, stable structure generated to date[5] and serves as a basis for **1** and **2**. Because of its simplicity, hub-(M)$_3$·3CA is a model system for evaluating different techniques useful for characterization, particularly those that determine the number and symmetries of isomers present in solution. As aggregates become more structurally complex, these methods of characterization become increasingly important.

Isomers of Aggregates 3. The imide region of the 1H NMR spectrum of **3** has proven to be diagnostic for the number and symmetries of aggregates in solution.[1] Figure 1 shows the two isomers of **3**. Isomerism is the result of the conformation of the linker arm (indicated with a ●). The imide protons are sensitive to the orientation of this connection, and at low temperatures, they are resolved in the NMR spectrum. The C_3 isomer,

(1) Simanek, E. E.; Wazeer, M. I. M.; Mathias, J. P.; Whitesides *J. Org. Chem.* **1994,** *59,* 4904.

(2) We have previously introduced DP as a measure of the relative stabilities of aggregates based on CA.M that could be easily obtained from computations. Briefly, the value of DP measures the extent that the CA.M groups are out of the mean plane of the CA.M rosette; the greater the deviation from planarity, the more unstable the aggregate becomes. For details, see refs 3 and 4.

(3) Chin, D. N.; Gordon, D. M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994,** *116,* 12033.

(4) Li, X.; Chin, D. N.; Whitesides, G. M. *J. Org. Chem.* **1996,** *61,* 1779.

Scheme 1. Aggregates Based on the CA·M Lattice

3a, gives two lines (the imide protons are numbered 1 and 2), and the C_1 isomer, **3b,** gives six lines (protons 3–8). The ratio of **3a:3b** depends weakly on solvent and temperature. We observe a 3:1 preference for trismelamine hubs in isomers with C_3 symmetry relative to those that have C_1 symmetry using a variety of different molecules of CA.

Isomers of 1 and 2. The increased complexity of **1** and **2** (five particles instead of four, the increased number of different types of protons, and the greater potential for isomerism) can lead to the formation of many isomers. As a result of this complexity, aggregates **1** and **2** offer
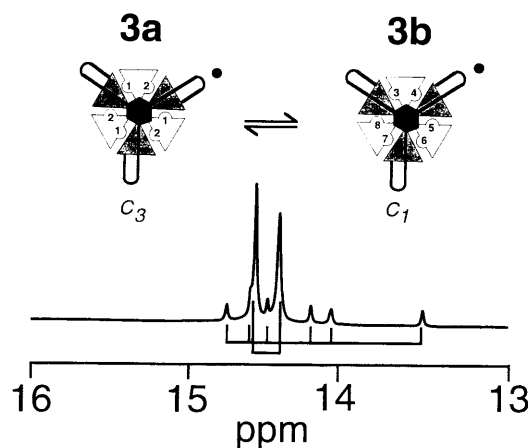

Figure 1. The Isomers of 3. The imide protons of two isomers of **2** produce two lines (C_3) and six lines (C_1). The unique environments of the protons are indicated with numbers 1–8. The isomers can be interconverted by rotation of the tether indicated with a ·.

Table 1. Isomers of 1 and 2 Obtained by Combining C_3 and C_1 Symmetric Hubs Are Shown[a]

Number	Scheme	Symmetry	Lines	C_3 Hubs	E.I.
i		D_3	2	2	0
ii		C_{3h}	2	2	3
iii		C_h	6	0	3
iv		C_2	6	0	2
v		C_2	6	0	2
vi		C_2	6	0	0
vii		C_1	12	0	1
viii		C_1	12	1	1
ix		C_1	12	1	2

[a] "Number" is used as a label. "Scheme" shows the relative orientations of the top hub (black lines) and bottom hub (white lines). "Symmetry" describes the symmetry of the resulting isomer. "Lines" shows the number of unique imide protons in the isomer: isomers showing similar numbers of lines are grouped with a dashed line. "C_3 hubs" shows the number of C_3 hubs in the aggregate; it refers only to the symmetry of the closest rosette. "E.I." shows the number of eclipsing interactions between anthranilate groups—judged unfavorable by computation—in the isomer. Stability is most favored in isomers with higher numbers of C_3 hubs and lower numbers of E.I.

an opportunity to test the generality of the rules established for the interpretation of the imide region and to determine the limitations of computation in assigning and predicting structure.[6]

Table 1 shows the isomers available to 1 and 2 from the different orientations and symmetries of the top and bottom hub(M)$_3$ groups. The number of different imide lines is dependent on the overall symmetry of the isomer. The observation that only one highly symmetric isomer (two lines) was present for 1 suggested an opportunity to use computation for assigning the structure of the observed isomer and determining the causes for the difference in stability.

Results

Aggregate 1 Exists as Mixture of Two Isomers. Figure 2 shows the imide region of the ¹H NMR spectra of 1 in chloroform recorded at 258 K. Fourteen lines are

(5) Whitesides, G. M.; Simanek, E. E.; Mathias, J. P.; Seto, C. T.; Chin, D. N.; Mammen, M.; Gordon, D. M. *Acc. Chem. Res.* **1995**, *37*.

(6) Crystal structures of 1–3 are not available. While we have been able to obtain a crystal structure of an aggregate based on the CA·M lattice no diffraction-quality crystals of 1–3 have been obtained. For crystal structures of tapes and of a discrete rosette, see: Zerkowski, J. A.; MacDonald, J. C.; Seto, C. T.; Wierda, D. A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 2382 and references therein.
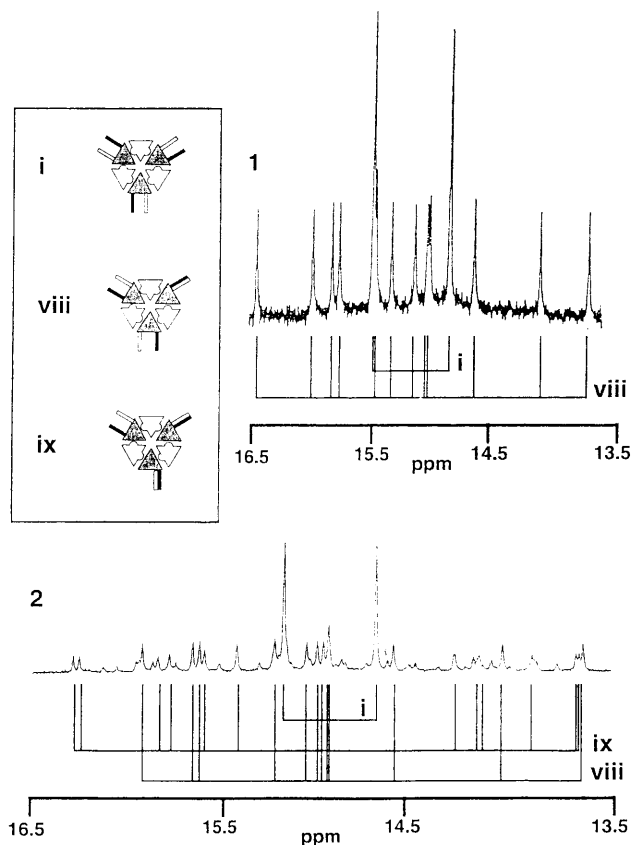

Figure 2. ¹H NMR of 1 and 2. The imide lines reflect the number and symmetries of isomers of 1 and 2. The lines corresponding to a single isomer are connected with the indicated trees. Aggregate 1 shows (two lines and 12 lines). Three isomers of aggregate 2 are indicated (2 lines, 12 lines, and 12 lines). Other isomers are present, but the similar sizes of the lines makes unambiguous assignments impossible. The relevant isomers from Table 1 are shown.

visible with small shifts occurring over the temperature range available to the solvent (mp–bp). The size of the lines suggests that two isomers (a and b) are present: one showing two lines and the second showing 12 lines. The intensities of these two sets of lines (two or 12) do not change over the temperature range surveyed (from the bp to mp of the solvent). The choice of solvent (CD$_2$-Cl$_2$, CDCl$_3$, C$_6$D$_6$) does not affect the relative intensities of lines in a set. While constant relative intensities of a series of the lines is not unambiguous proof of one isomer (of 12 lines) instead of two isomers (of six lines) in identical concentrations, it is consistent with trends in assignment that will be discussed.

Aggregate 2 Exists as a Mixture of at Least Four Isomers. Figure 2 shows the imide region for 2 recorded at 278 K. From the intensities of the lines, we believe that at least four aggregates are present. We refer to these aggregates as 2a–2c. Additional lines that are smaller than those of 2c are observed. It is difficult to classify these into sets of 12 lines or 6 lines.

Computations Suggests That Isomer i Is More Stable Than Isomer ii. From ¹H NMR spectroscopy we know that only one of two high-symmetry isomers (i or ii) is present in solution. We wanted to see if the computational surrogate for stability, DP, could be used to distinguish these isomers and to determine the causes for the difference in stability between i and ii.[3]

To be of general use to organic chemists, computational methods should be efficient and not require extensive

Scheme 2. Construction of the Computational Model[a]
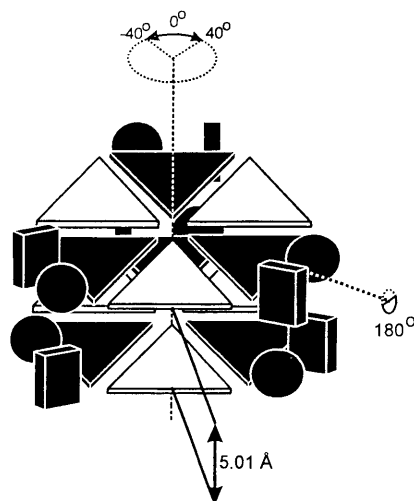
[a] The initial distance between rosettes is set to be 5.01 Å—the distance between CH_2 groups of the benzCA$_2$ linker group.

Figure 3. The values of DP for the various arrangements of rosettes (Figure 3) in the two isomers. Only those data where distance between the methylene groups is 5.01 Å is shown.

amounts of time. This requirement excludes free-energy perturbation methods on molecules as complex as these.[7–9] Our approach in determining relative stabilities between the isomers, therefore, involves a more empirical strategy: we include only the primary groups in contact—the rosettes—in the computations. This approach is based on results from previous work,[3,4] which suggested that groups in contact with the rosette determine the stability of the aggregate. No solvent is included in the model.

In this simplified model, the molecules are melamines that have phenyl and neohexyl substituents and isocyanurates with methyl substituents—the presence of the *m*-xylyl linker is implicitly included using constraints (as we will discuss). An atomistic model of these molecules was constructed using standard valence geometries in the Quanta 3.3 molecular modeling program,[10] and CHARMm 22 was used for molecular mechanics calculations.[11] In these calculations, no cutoff values are used in determining nonbonded interactions, and the dielectric constant is set to 1. The molecules in each of the two rosettes were placed into a coplanar arrangement separated by a hydrogen-bond distance of 1.8 Å.[12] Each isomer is modeled in the following ways. Isomer **i** is assembled by separating the two rosettes by a distance of 5.01 Å (the distance between methylene groups of the *m*-xylyl linker) between the planes defined by the atoms in the rosette. This initial arrangement is such that the two rosettes are related by mirror symmetry about a plane located equidistant between the rosettes and parallel to the rosettes (Scheme 2). Isomer **ii** is assembled similarly to isomer **i** except that the three melamine groups in one of the rosettes are rotated 180° about an axis defined by the N atom in the $-NH_2$ group and the N atom at the four position.

Starting from the initial arrangement of the two rosettes for each isomer, one rosette is rotated ±40°, in

1° increments, relative to the other rosette about an axis that is normal to the rosettes and located at their centers (Scheme 2). At each iteration in the rotation, the energy of the system is minimized using the Adopted-Basis Newton Raphson (ABNR) method[11] until the gradient in the energy is less than 0.1 kcal/mol·Å. During the minimizations, constraints were applied to the system in two areas. First, the distance between the methyl groups of proximate cyanurates were fixed to a value of 5.01 to mimic the constraint placed on these groups by the *m*-xylyl linker group. Second, the conformation of the phenyl substituents on the melamines was held perpendicular (90°) to the plane of the melamines to be consistent with previous results from molecular dynamics simulations of Hub(M)$_3$·3CA.[3] For completeness, the computations were carried out within a ±40° window.

Discussion

As an alternative measure of relative stabilities between complexes, we have used the deviation from planarity (DP) for atoms in the rosettes for comparing the relative stabilities of various complexes based on CA·M. For these complexes, lower values of DP—the constituent molecules become more coplanar with each other—are associated with increasing stabilities.[3] Relative values of DP from the various arrangements of the two isomers in the computations suggest that isomer **i** is also favored over isomer **ii** (Figure 3). The data in Figure 3 suggest two trends: the favorable (low) values of DP for isomer **i** plateau between +10° and +30° and there is a rotational preference to the arrangement of the rosettes. An example of the structure of isomer **i** from this favorable region is shown in Figure 4. The relative values of DP for these two isomers, consequently, are consistent with the experimental observation that isomer **i** is more stable than isomer **ii**.

The values of DP obtained for **1** are consistent with the observed experimental stability of the aggregate: we have established that **1** is more stable than **3** using gel permeation chromatography and titrations of these aggregates with methanol. The value of DP for **3** is DP = 0.3 Å; that for **1** is 0.2 Å.[3]

Three Criteria Determine the Assignment of Isomers of 1 and 2: Number of Imide Lines, Symmetries of Hub Groups, and Relative Orientations of Hub Groups as Suggested by coMputation. The three criteria used in assigning the structures of ag-

(7) Beveridge, D. L.; Dicapua, F. M. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431.

(8) Mitchell, M. J.; McCammon, A. J. *J. Comput. Chem.* **1991**, *12*, 271.

(9) Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 3770.

(10) *Quanta 3.3 Parameter Handbook.* MSI: Burlington, MA, 1992.

(11) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. L. *J. Comput. Chem.* **1983**, *4*, 187.

(12) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810.
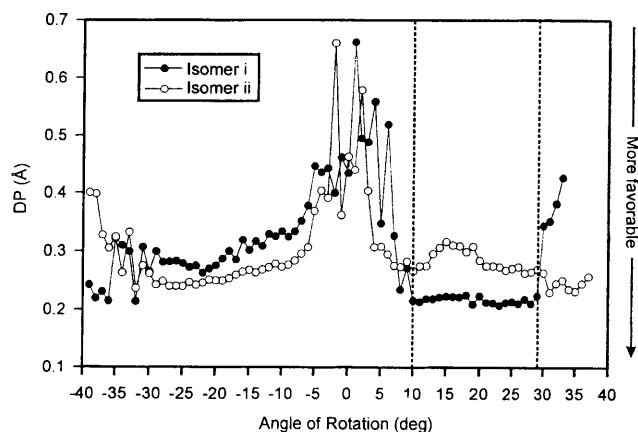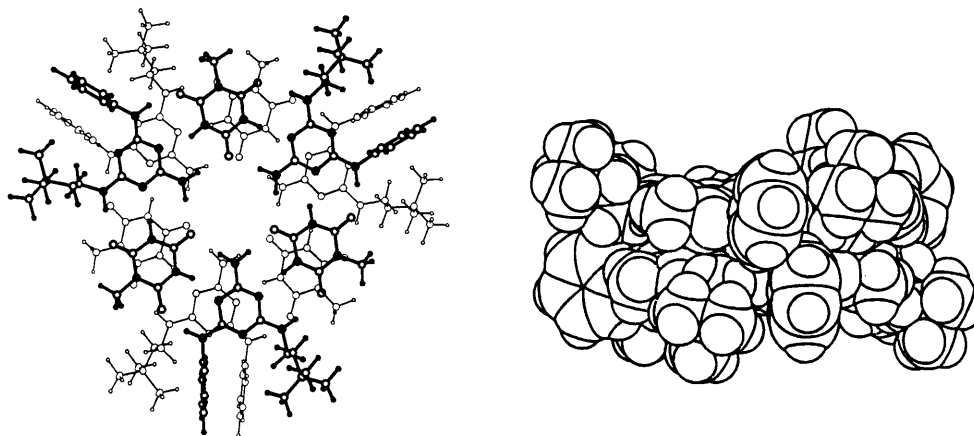
Figure 4. Top view and side views of isomer **i** after minimization of the conformer obtained after rotating 23°. This structure has the lowest value of DP.

gregates **1** and **2** are detailed below. We conclude that the isomers **1a** and **2a** are **i** and that isomers **1b** and **2b** are **viii**. Isomer **2c** is **ix**.

Number of Imide Lines. The most important clue to the identity of an isomer is the number of imide lines in the ^{1}H NMR spectrum. Isomers with 2, 6, or 12 lines are the only possibilities. The number of lines for each isomer is shown on Table 1.

Number of Symmetric Hubs. Aggregate **3** exists as a mixture of C_3 and C_1 isomers. The C_3 isomer is favored by at least 0.5 kcal in all systems studied.[1] This value corresponds roughly to a 3:1 preference for the C_3 isomer over the C_1 isomer. If this preference is the result of favorable interactions (or lack of unfavorable interactions) resulting from the arrangements of the arms of the hub(M)$_3$ molecule when complexed with molecules of CA, then it is reasonable to expect that this preference will exist in aggregates **1** and **2**, where it is complexed to bisCA. Consistent with this belief is the observation that the isomers producing two lines must have two C_3 hub-(M)$_3$ groups. Table 3 shows the number of hub(M)$_3$ groups with C_3 and C_1 orientations for each isomer. Of the three isomers giving 12 lines, two incorporate C_3 hub-(M)$_3$ groups. The reasons for the preference for symmetric hub groups is uncertain.

Orientation of the Hubs. The CA rings of the bisCA groups are constrained to lie closely in registry. The melamine rings, therefore, must also lie in approximate registry. The peripheral groups of the melamine (either an anthranilate or neohexyl group) must be located close to each other. The fact that we see only one isomer with two lines (instead of two) suggests that there may be a strong preference for the relative orientations of the anthranilates and neohexyl groups. We call the placement of two groups (anthranilates) in approximate registry an *eclipsing* interaction. In isomer **i** (two lines, two C_3 hubs) no anthranilate/anthranilate eclipsing interactions take place. In isomer **ii** (two lines, two C_3 hubs) two anthranilate/anthranilate eclipsing interactions occur. We had no method of determining whether these types of interactions were favorable or unfavorable (that is, whether **i** or **ii** was present in **1**) before we modeled these interactions.

Assignment of Isomers. The three criteria described above can be used in assigning the structures of aggregates **1** and **2**. Starting with the number of imide lines, we conclude that the isomers **1a** and **2a** are **i** or **ii**. We also conclude that isomers **2b,c** are **viii**, **vii**, and/or

ix. Complete assignment of the isomers of **1** and **2** can be made by assuming that isomers that have fewer anthranilates in register (lower values of E.I.) will be more stable than those with many eclipsed anthranilate groups: **i** is more stable than **ii**. We predict that the isomer showing two lines in **1a** and **2a** is **i** (two C_3 hubs, zero E.I.) and not **ii** (two C_3 hubs, three E.I.). Isomer **1b** and **2b** are likely to be **viii** (one C_3 hub, one E.I.) over **iv** (one C_3 hub, two E.I.) or **vii** (zero C_3 hub, one E.I.). To assign isomer **2c**, we must make a judgment based on which is worse: the exclusion of C_3 hubs or large numbers of E.I. Based on the observation of isomers with 12 lines and none with six lines it is likely that C_3 hubs are more prevalent than C_1 hubs. We tentatively assign **2c** as **ix**.

Conclusions

Using computation, we have tested general rules for the assignment of isomers of aggregates **1** and **2**. *The most interesting issue arising from this work is that a simple modification to a noncovalent aggregate can translate into additional structural simplicity.* That is, the use of iPr-benzCA$_2$ instead of iPr-furanCA$_2$ results in fewer isomers. The rules for the interpretation of the imide region of the ^{1}H NMR spectrum apply to these more complicated aggregates: they suggest the number and symmetries of isomers in solution. Additional proof-of-structure for aggregates **1** and **2** is, as a result, available. Solvent and temperature appears to have minimal effect on the number and symmetries of these aggregates. Aggregates incorporating C_3 symmetric hub-(M)$_3$ groups are more stable than those that incorporate asymmetric hub(M)$_3$ groups. Aggregate **2**, with a bisCA$_2$ of geometry close to ideal, provides for much greater conformational isomerism than the less-than-ideal benzCA$_2$ of **1**. Learning to characterize and control differences in structure during the preparation of noncovalent aggregates remains a goal for us and the community involved in molecular self-assembly. The increased generality of DP—to aggregates comprising stacked rosettes—is a small step toward this goal.

Experimental Section

Synthesis of 1–3. The synthesis and characterization of these aggregates has been reported.[13,14]

(13) Syntheses of **1** and **2**: Seto, C. T.; Mathias, J. P.; Whitesides, G. M. *J. Am. Chem. Soc.* **1993**, *115*, 1321.
(14) Synthesis of **3**: Seto, C. T.; G. M., W. *J. Am. Chem. Soc.* **1993**, *115*, 1330.

Sample Preparation. Deuterated solvent obtained from Aldrich (CD_2Cl_2) and the Cambridge Isotope Laboratories (o-$C_6D_4Cl_4$ and $C_2D_2Cl_4$) were used as received. Chloroform-d was Aldrich grade. Spectra were recorded at 10 mM in 0.5 mL of solvent.

Variable Temperature Spectra. The data were recorded on Bruker AM-400 spectrometer. FIDs (32 or 64 scans) were collected using a 90° pulse width (14.5 sec) with a short recycling delay[15] (1 s), and an exponential correction (0.2 Hz line broadening) was applied. Temperatures are accurate to within one degree as judged by a Wescon 640D unit. Indistinguishable spectra were generated at a particular temperature on raising and lowering the temperature.

---

(15) Varying the recycling delays from 1 to 30 s did not affect the relative sizes of the lines in the imide region. We selected a short recycling delay to expedite experiments and in order to minimize the variations in temperature occurring during long experiments.