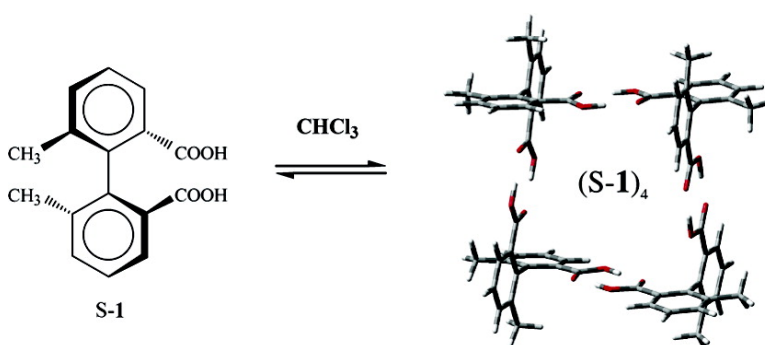


Determination of Molecular Structure in Solution Using Vibrational Circular Dichroism Spectroscopy: the Supramolecular Tetramer of *S*-2,2'-Dimethyl-Biphenyl-6,6'-Dicarboxylic Acid

M. Urbanov, V. Setnic#ka, F. J. Devlin, and P. J. Stephens

J. Am. Chem. Soc., **2005**, 127 (18), 6700-6711 • DOI: 10.1021/ja050483c • Publication Date (Web): 19 April 2005

Downloaded from <http://pubs.acs.org> on March 25, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 10 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)

Determination of Molecular Structure in Solution Using Vibrational Circular Dichroism Spectroscopy: the Supramolecular Tetramer of S-2,2'-Dimethyl-Biphenyl-6,6'-Dicarboxylic Acid

M. Urbanová,[†] V. Setnička,[†] F. J. Devlin,[‡] and P. J. Stephens^{*‡}

Contribution from the Department of Physics and Measurements and Department of Analytical Chemistry, Institute of Chemical Technology, Technická 5, Prague 6, 166 28, Czech Republic, and Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482

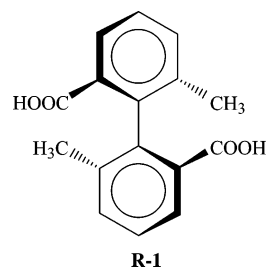
Received January 24, 2005; E-mail: pstephen@usc.edu

Abstract: The infrared (IR) and vibrational circular dichroism (VCD) spectra of S-2,2'-dimethyl-biphenyl-6,6'-dicarboxylic acid, **S-1**, in CDCl₃ solution are concentration-dependent, showing that oligomerization occurs with increasing concentration. DFT calculations support the conclusion that the oligomer formed is the cyclic tetramer (**S-1**)₄, in which **S-1** monomers are linked by hydrogen(H)-bonded (COOH)₂ moieties. Due to the existence of two inequivalent tautomeric conformations of each (COOH)₂ moiety, six inequivalent conformations of (**S-1**)₄ are possible. B3LYP/6-31G* DFT calculations predict that the conformation "aaab", possessing three equivalent (COOH)₂ conformations, a, and one tautomeric conformation, b, has the lowest free energy. B3LYP/6-31G* IR and VCD spectra vary substantially with conformation. The B3LYP/6-31G* IR and VCD spectra of the C=O stretch modes of "aaab" are in excellent agreement with the experimental spectra, while those of all other conformations exhibit poor agreement, confirming the prediction that the "aaab" conformation is the predominant conformation. Comparison of the calculated IR and VCD spectra of the six conformations to the experimental spectra in the range 1100–1600 cm⁻¹ further supports this conclusion. The study is the first to use VCD spectroscopy to determine the structure of a supramolecular species.

Introduction

X-ray crystallography by Tichý et al.¹ has shown that crystals of the R enantiomer of the chiral molecule 2,2'-dimethyl-biphenyl-6,6'-dicarboxylic acid, **1**, contain intermolecularly hydrogen(H)-bonded cyclotetramers of **1**. Tichý et al. also reported that vapor phase osmometry indicated "an apparent degree of association of ~4" in CHCl₃ solution at concentrations higher than 0.15 mol·kg⁻¹. However, the aggregation of **1** in solution has not been further characterized, and the question as to whether self-assembly of **1** forming supramolecular cyclotetramers occurs in solution remains open. In view of the increasing importance of supramolecular chemistry,² we seek here to clarify this issue.

The infrared (IR) vibrational absorption spectrum of a molecule is a sensitive function of its structure and, in the case of molecules which aggregate, of supramolecular structure. Chiral molecules also exhibit vibrational circular dichroism (VCD).³ VCD spectra are even more sensitive to molecular structure than are IR spectra and, in the case of molecules which



aggregate, can be expected to be exquisitely sensitive to supramolecular structure. We have therefore undertaken a study of the IR and VCD spectra of solutions of **1**. Our primary goal is to determine whether a cyclic tetramer of **1** is indeed formed in CHCl₃ solution and, if so, to characterize its structure.

Achievement of this goal requires a methodology by which structural information can be extracted from IR and VCD spectra. Our methodology utilizes ab initio density functional

[†] Institute of Chemical Technology.

[‡] University of Southern California.

(1) Tichý, M.; Kraus, T.; Závada, J.; Císařová, I.; Podlaha, J. *Tetrahedron: Asymmetry* **1999**, *10*, 3277–3280.

(2) (a) Lehn, J.-M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4763–4768. (b) Lehn, J.-M. *Polym. Int.* **2002**, *51*, 825–839. (c) Lehn, J.-M. *Science* **2002**, *295*, 2400–2403. (d) Lehn, J.-M. *Rep. Prog. Phys.* **2004**, *67*, 249–265.

(3) (a) Stephens, P. J.; Lowe, M. A. *Annu. Rev. Phys. Chem.* **1985**, *36*, 213–241. (b) Stephens, P. J. *Encyclopedia of Spectroscopy and Spectrometry*; Academic Press: London, 2000; pp 2415–2421. (c) Stephens, P. J.; Devlin, F. J. *Chirality* **2000**, *12*, 172–179. (d) Stephens, P. J.; Devlin, F. J.; Aamouche, A. In *Chirality: Physical Chemistry*; Hicks, J. M., Ed., ACS Symposium Series 810; American Chemical Society: Washington, DC, 2002; Chapter 2, pp 18–33. (e) Stephens, P. J. In *Computational Medicinal Chemistry for Drug Discovery*; Bultinck, P., de Winter, H., Langenaecker, W., Tollenaere, J., Eds.; Dekker: New York, 2003; Chapter 26, pp 699–725.

theory (DFT).⁴ The reliability of DFT predictions of IR and VCD spectra is by now well-documented,⁵ and the methodology used here has been extensively utilized for the conformational and configurational analysis of chiral molecules.⁶ Our study of **1** extends prior work to the field of supramolecular chemistry.

Results

Experimental IR and VCD Spectra of 1. The IR and VCD spectra of 0.16 M CDCl₃ and 0.19 M DMSO-*d*₆ solutions of optically pure *S*-(+)-**1** ([α]_D +21.7° (*c* 1, MeOH)) in the mid-IR spectral region are shown in Figure 1. Substantially different spectra are observed in the two solvents. The difference is especially large between the VCD spectra. Overall, the VCD is much larger in CDCl₃ than in DMSO-*d*₆. The increase in intensity is most dramatic in the C=O stretching region (~1700 cm⁻¹). In DMSO-*d*₆, a very small bisignate feature is observed. In CDCl₃, by contrast, a very intense bisignate feature is observed: the ratios of peak Δε values to the peak ε are > 1 × 10⁻⁴.

The concentration dependence of the IR and VCD spectra of *S*-**1** in CDCl₃ solution in the mid-IR spectral region over the concentration range 1.8 mM to 0.16 M is shown in Figure 2. In the C=O stretching region, the bands in the IR spectrum at 1703 and 1683 cm⁻¹ increase in intensity with increasing concentration, while the band at 1738 cm⁻¹ decreases in intensity. In the VCD spectrum the bisignate feature increases dramatically in intensity with increasing concentration.

The observation of concentration-dependent IR and VCD spectra in CDCl₃ demonstrates that aggregation of *S*-**1** occurs in this solvent with increasing concentration and that the spectra at higher concentrations originate in one or more multimers of *S*-**1**. Dimerization of carboxylic acids in nonpolar solvents via intermolecular hydrogen(H)-bonding is well-known, of course, and aggregation of **1** is not surprising. However, given the presence of two COOH groups in **1**, multiple multimeric forms are possible. One of these is the cyclic tetramer identified in crystalline *R*-**1** by X-ray crystallography,¹ in which all COOH

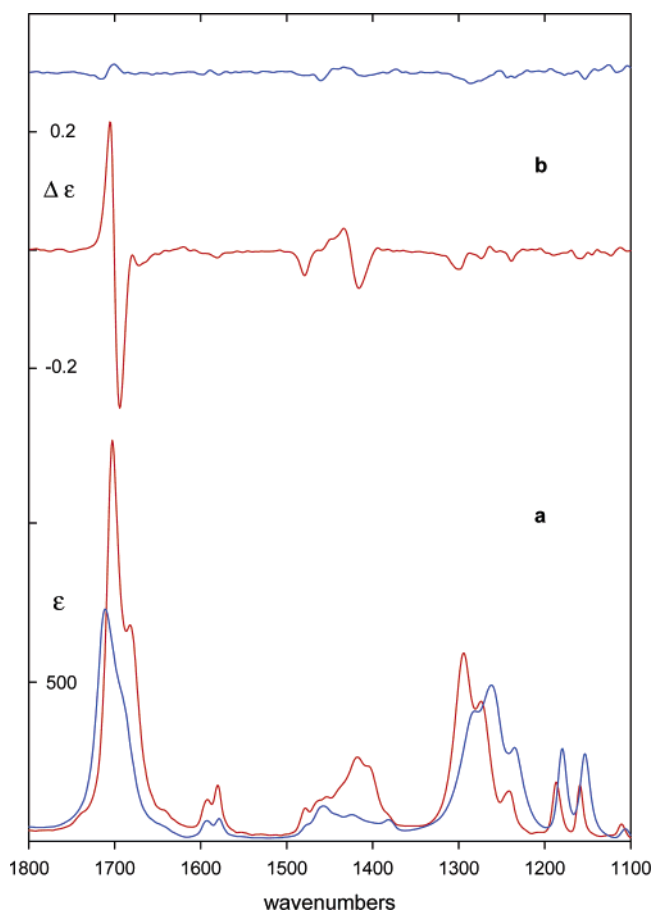
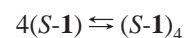


Figure 1. IR (a) and VCD (b) spectra of *S*-(+)-**1** in CDCl₃ (red, 0.16 M) and DMSO-*d*₆ (blue, 0.19 M). Path length 50 μm.

groups are involved in intermolecularly H-bonded dimeric (COOH)₂ structures. It is reasonable to hypothesize that the aggregation of *S*-**1** in CDCl₃ solution involves tetramerization:



and that the IR and VCD spectra at the highest concentrations originate in the cyclic tetramer of *S*-**1**. The primary goal of this paper is to establish whether this hypothesis is indeed correct.

The concentration dependence of the IR spectrum of CDCl₃ solutions of *S*-**1** in the O–H stretching region provides additional information regarding the aggregation of **1**. Monomeric **1** contains two “free” (i.e., non-H-bonded) COOH groups which are expected to give rise to IR absorption in the O–H stretching region. H-bonding, forming (COOH)₂ moieties, is expected to substantially decrease O–H stretching frequencies. Formation of a cyclic tetramer of **1** in which no COOH groups are “free” would then be expected to be accompanied by complete disappearance of the O–H stretching absorption of the “free” COOH groups of monomeric **1**. The IR spectra of 0.16 M, 19 mM, 5 mM, 3.6 mM, and 1.8 mM solutions of *S*-**1** in CDCl₃ in the hydrogenic stretching region are shown in Figure 3. A band at 3512 cm⁻¹ is observed at 1.8 mM which decreases in intensity with increasing concentration and is barely visible at 0.16 M. We assign this band to the O–H stretch of the “free” COOH groups of monomeric *S*-**1**. Its disappearance at high concentration is consistent with the formation of a cyclic tetramer containing only H-bonded COOH groups, whose O–H stretching modes are shifted substantially to lower frequencies.

- (4) Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. *J. Chem. Phys. Lett.* **1996**, *252*, 211–220. (b) Stephens, P. J.; Ashvar, C. S.; Devlin, F. J.; Cheeseman, J. R.; Frisch, M. J. *Mol. Phys.* **1996**, *89*, 579–594.
- (5) (a) Devlin, F. J.; Stephens, P. J.; Cheeseman, J. R.; Frisch, M. J. *J. Phys. Chem.* **1997**, *101*, 6322–6333. (b) Devlin, F. J.; Stephens, P. J.; Cheeseman, J. R.; Frisch, M. J. *J. Phys. Chem.* **1997**, *101*, 9912–9924. (c) Ashvar, C. S.; Devlin, F. J.; Stephens, P. J.; Bak, K. L.; Eggimann, T.; Wieser, H. *J. Phys. Chem. A* **1998**, *102*, 6842–6857.
- (6) (a) Ashvar, C. S.; Devlin, F. J.; Stephens, P. J. *J. Am. Chem. Soc.* **1999**, *121*, 2836–2849. (b) Devlin, F. J.; Stephens, P. J. *J. Am. Chem. Soc.* **1999**, *121*, 7413–7414. (c) Aamouche, A.; Devlin, F. J.; Stephens, P. J. *J. Am. Chem. Soc.* **2000**, *122*, 2346–2354. (d) Aamouche, A.; Devlin, F. J.; Stephens, P. J. *J. Am. Chem. Soc.* **2000**, *122*, 7358–7367. (e) Aamouche, A.; Devlin, F. J.; Stephens, P. J.; Drabowicz, J.; Bujnicki, B.; Mikolajczyk, M. *Chem.–Eur. J.* **2000**, *6*, 4479–4486. (f) Stephens, P. J.; Aamouche, A.; Devlin, F. J.; Superchi, S.; Donnoli, M. I.; Rosini, C. *J. Org. Chem.* **2001**, *66*, 3671–3677. (g) Devlin, F. J.; Stephens, P. J.; Scafato, P.; Superchi, S.; Rosini, C. *Tetrahedron: Asymmetry* **2001**, *12*, 1551–1558. (h) Devlin, F. J.; Stephens, P. J.; Scafato, P.; Superchi, S.; Rosini, C. *Chirality* **2002**, *14*, 400–406. (i) Devlin, F. J.; Stephens, P. J.; Oesterle, C.; Wiberg, K. B.; Cheeseman, J. R.; Frisch, M. J. *J. Org. Chem.* **2002**, *67*, 8090–8096. (j) Devlin, F. J.; Stephens, P. J.; Scafato, P.; Superchi, S.; Rosini, C. *J. Phys. Chem. A* **2002**, *106*, 10510–10524. (k) Ceré, V.; Peri, F.; Pollicino, S.; Ricci, A.; Devlin, F. J.; Stephens, P. J.; Gasparrini, F.; Rompietti, R.; Villani, C. *J. Org. Chem.* **2005**, *70*, 664–669. (l) Setnička, V.; Urbanová, M.; Bouř, P.; Král, V.; Volka, K. *J. Phys. Chem. A* **2001**, *105*, 8931–8938. (m) Bouř, P.; Navrátilová, H.; Setnička, V.; Urbanová, M.; Volka, K. *J. Org. Chem.* **2002**, *67*, 161–168. (n) Butkus, E.; Zilinskas, A.; Stoncius, S.; Rozenbergas, R.; Urbanová, M.; Setnička, V.; Bouř, P.; Volka, K. *Tetrahedron: Asymmetry* **2002**, *13*, 633–638. (o) Urbanová, M.; Setnička, V.; Bouř, P.; Navrátilová, H.; Volka, K. *Biopolymers (Biospectroscopy)* **2002**, *61*, 298–301. (p) Tománková, Z.; Setnička, V.; Urbanová, M.; Matějka, P.; Král, V.; Volka, K.; Bouř, P. *J. Org. Chem.* **2004**, *69*, 26–32.

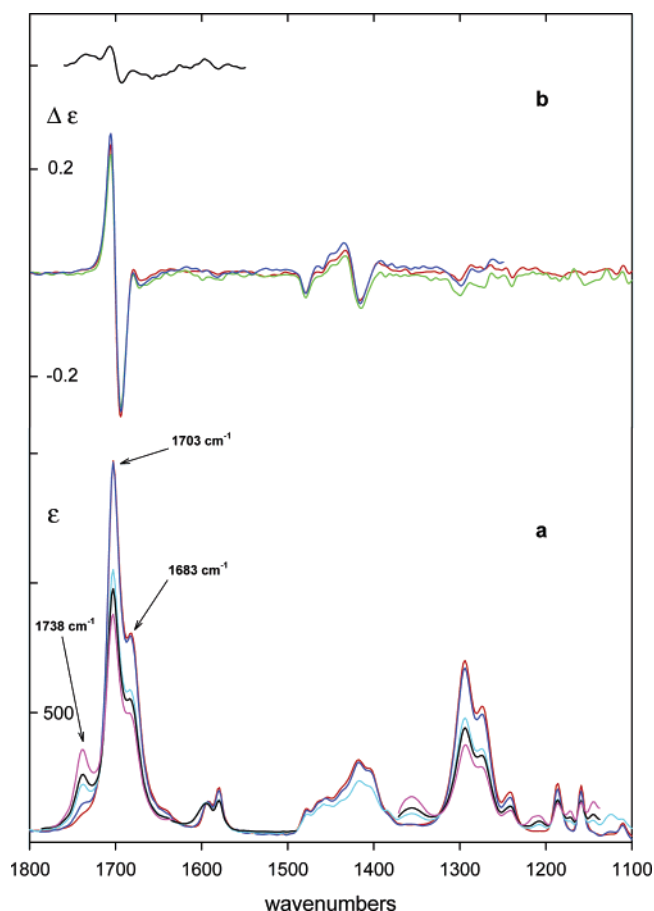


Figure 2. Concentration dependence of the IR (a) and VCD (b) of *S*-(+)-**1** in CDCl_3 solution in the mid-IR spectral region. Concentrations/path lengths are: IR: 0.16 M/50 μm (red), 0.019 M/210 μm (blue), 0.005 M/545 μm (cyan), 0.0036 M/1.0 mm (black), 0.0018 M/1.0 mm (pink). VCD: 0.16 M/50 μm (red), 0.08 M/50 μm (green), 0.019 M/210 μm (blue), 0.0036 M/1.0 mm (black).

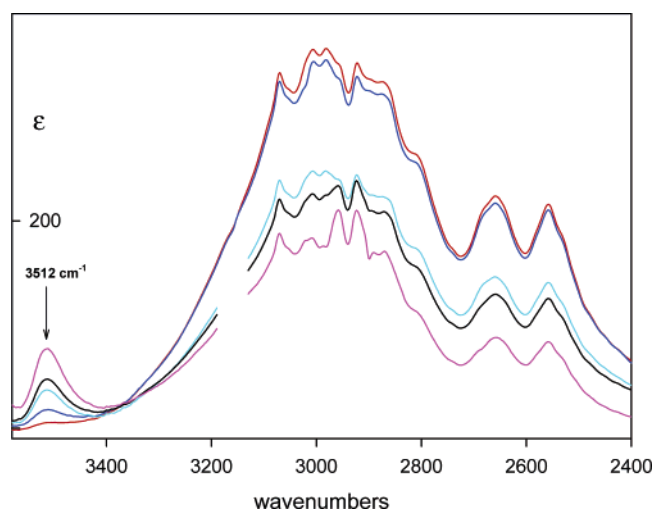


Figure 3. Concentration dependence of the IR of *S*-(+)-**1** in CDCl_3 solution in the hydrogenic stretching region. Concentrations/path lengths are: 0.16 M/50 μm (red), 0.019 M/210 μm (blue), 0.005 M/545 μm (cyan), 0.0036 M/1.0 mm (black), and 0.0018 M/1.0 mm (pink). The gap is due to excessive solvent absorption.

To further analyze the IR and VCD spectra of CDCl_3 solutions of *S*-**1**, we have carried out DFT calculations of the structure(s) and IR and VCD spectra of the cyclic tetramer(s) of *S*-**1**. Before describing the results of these calculations, we

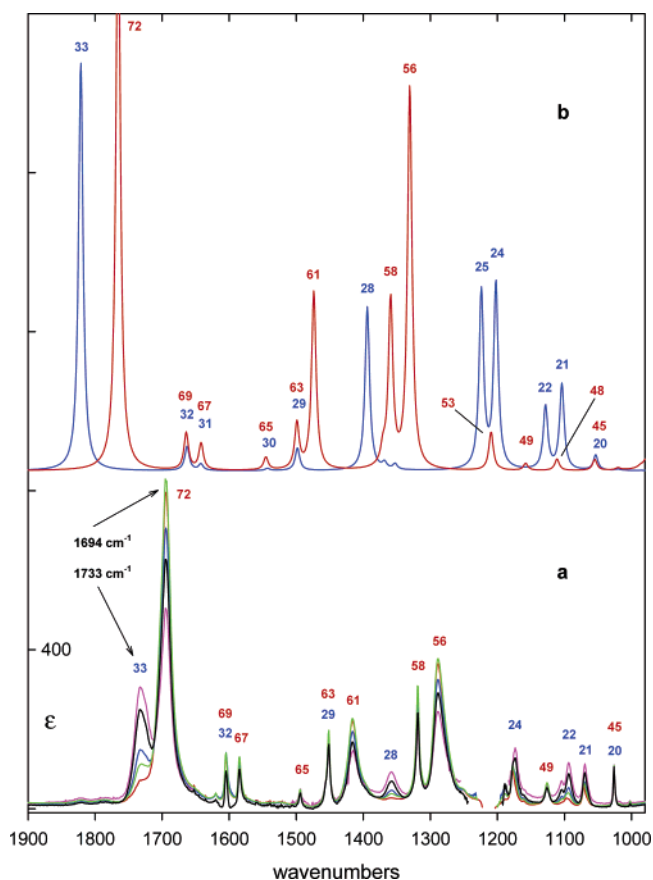
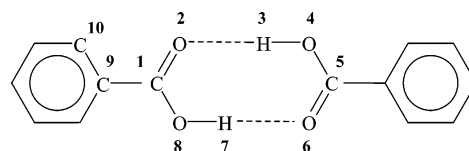


Figure 4. (a) Experimental IR spectra of benzoic acid, **2**, in CHCl_3 solution in the mid-IR spectral region. Concentrations/path lengths are: 0.37 M/45 μm (red), 0.093 M/216 μm (green), 0.037 M/215 μm (blue), 0.0093 M/500 μm (black), 0.0047 M/1.0 mm (pink). Assignments to modes of monomeric and dimeric forms of **2**, based on the calculated spectra in (b), are shown. (b) Calculated, B3LYP/6-31G*, IR spectra of the monomeric (blue) and dimeric (red) forms of **2** in the mid-IR spectral region. Fundamentals are numbered. Lorentzian bandwidth parameter $\gamma = 4.0 \text{ cm}^{-1}$.

make a brief detour to discuss the IR spectra of benzoic acid, **2**, in CHCl_3 solution.

IR Spectrum of Benzoic Acid. The dimer of benzoic acid, (**2**)₂, formed by intermolecular H-bonding of two carboxyl groups of two molecules of **2**:



constitutes an excellent model for the intermolecular interaction of the carboxyl groups of **1**. In order to calibrate our expectations with regard to the structures and vibrational spectra of (*S*-**1**)₄, we have examined the concentration dependence of the IR spectrum of **2** in CHCl_3 solution and the structures and IR spectra predicted using DFT for the monomeric and dimeric forms of **2**.

The concentration dependence of the IR spectrum of **2** in CHCl_3 solution over the concentration range 0.004–0.4 M in the mid-IR and O–H stretching regions of the spectrum is shown in Figures 4 and 5. The spectrum is strongly concentration-dependent. In the C=O stretching region, two bands are observed at 1733 and 1694 cm^{-1} whose intensities decrease and

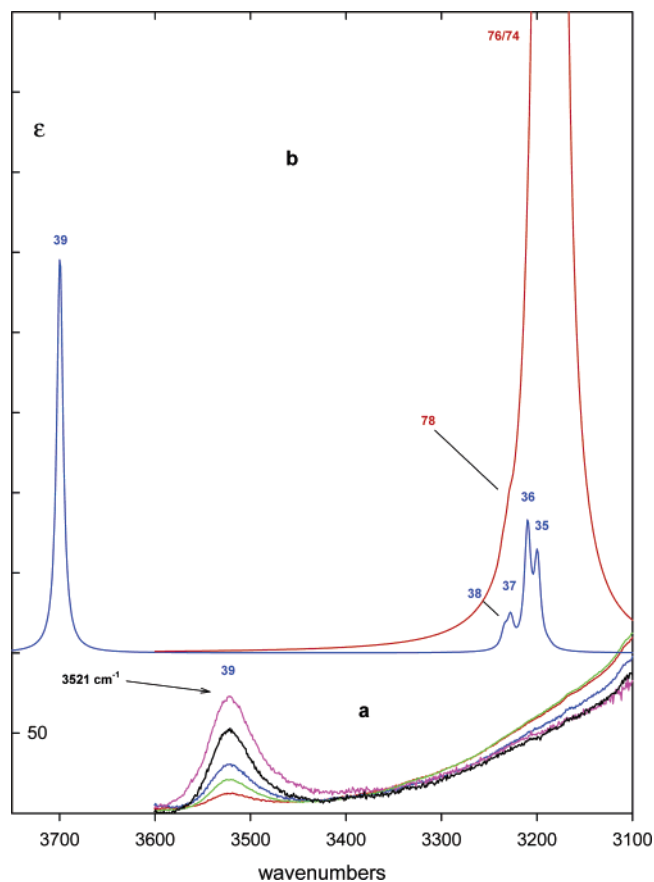


Figure 5. (a) Experimental IR spectra of benzoic acid, **2**, in CHCl_3 solution in the O–H stretching region. Concentrations/path lengths are: 0.37 M/45 μm (red), 0.093 M/216 μm (green), 0.037 M/215 μm (blue), 0.0093 M/500 μm (black), 0.0047 M/1.0 mm (pink). (b) Calculated, B3LYP/6-31G*, IR spectra of the monomeric (blue) and dimeric (red) forms of **2** in the O–H stretching region. Fundamentals are numbered. Lorentzian bandwidth parameter $\gamma = 4.0 \text{ cm}^{-1}$.

increase, respectively, with increasing concentration. An isosbestic point is observed at 1712 cm^{-1} , demonstrating that the spectra reflect an equilibrium between two species only. In the O–H stretching region, a band is observed at 3521 cm^{-1} whose intensity decreases with increasing concentration. We assign the bands at 1733 and 3521 cm^{-1} to monomeric **2** and the band at 1694 cm^{-1} to dimeric **2**, (**2**)₂. On this basis, it is clear that at the highest concentration, 0.37 M, the IR spectrum is predominantly due to (**2**)₂, while at the lowest concentration, 4.7 mM, the spectrum reflects a mixture of **2** and (**2**)₂. Experiments at substantially lower concentrations would be required to obtain the IR spectrum of **2**, unmixed with that of (**2**)₂.

Support for our analysis of the IR spectrum of **2** and a more detailed assignment of the mid-IR spectrum is provided by DFT calculations. The B3LYP/6-31G* harmonic vibrational frequencies and dipole strengths of **2** and (**2**)₂ are given in Table 1 of the Supporting Information (SI). The IR spectra obtained thence in the mid-IR and O–H stretching regions are shown in Figures 4 and 5, respectively. In (**2**)₂, there are two C=O stretching modes, symmetric (s) and anti-symmetric (as), at 1721 and 1766 cm^{-1} , respectively. Due to the C_{2h} symmetry of (**2**)₂, only the as mode is IR-allowed. Its predicted frequency is 55 cm^{-1} lower than the predicted C=O stretching frequency of **2**, 1821 cm^{-1} . The predicted spectra support the assignment of the 1733 and 1694 cm^{-1} bands in the experimental spectrum to **2** and (**2**)₂,

the predicted frequency shift of 55 cm^{-1} being commensurate with the experimental shift of 39 cm^{-1} . The assignment of the IR spectrum of **2** over the range 1000 – 1650 cm^{-1} is indicated in Figure 4. Bands assigned to **2** and (**2**)₂ alone decrease and increase in intensity, respectively, with increasing concentration. The qualitative agreement between theory and experiment is excellent. The O–H stretching frequencies of **2** and (**2**)₂ are predicted at 3700 and $3188/3185 \text{ cm}^{-1}$, respectively. The large shift to lower frequency on dimerization is consistent with experiment (Figure 5): the O–H stretch IR of (**2**)₂ is buried within the broad, complex absorption originating in the C–H stretching modes.

We have examined the basis set dependence of the B3LYP IR spectrum of (**2**)₂, increasing the basis set to 6-311G** and 6-311++G**, with the results shown in Figure 1 of the SI (see Table 1 of the SI for the harmonic frequencies and dipole strengths). Qualitatively, the spectra are very insensitive to the increases in basis set size. Quantitatively, the spectra exhibit an overall shift to lower frequency with increasing basis set size and, as a result, improved agreement of calculated and experimental frequencies.

Having validated the reliability of the DFT IR spectrum of (**2**)₂, it is of interest to examine the predicted structure of (**2**)₂ and, especially, the changes from the structure of **2**. Key structural parameters of **2** and (**2**)₂ are given in Table 2 of the SI. At the 6-31G* basis set level, from **2** to (**2**)₂ the C=O bond lengthens by 0.02 \AA , the C–O bond shortens by 0.035 \AA , and the O–H bond lengthens by 0.03 \AA . The O···H and O···O H-bonding distances in (**2**)₂ are 1.673 and 2.677 \AA . The O···H–O angle is 178.6° . Increasing the basis set size to 6-311G** and 6-311++G** leads to only minor changes in the structural parameters of (**2**)₂.

The last issue of interest here is the barrier to the switching of the two H-bonded H atoms of the (COOH)₂ moiety between their two possible configurations. We have scanned the potential energy surface (PES) of (**2**)₂ with respect to the positions of these two H atoms at the B3LYP/6-31G* level with the results shown in Figure 6. The two minima observed correspond to the two equivalent tautomers of the (COOH)₂ moiety of (**2**)₂. The barrier to interconversion between these two tautomers is predicted to be 7 – 8 kcal/mol .

Conformational Analysis of the Cyclic Tetramer of *S*-1.

The cyclic tetramer of *S*-1, (*S*-1)₄, contains four (COOH)₂ moieties. As the above discussion of the benzoic acid dimer has illustrated, there are two stable tautomeric conformations of each (COOH)₂ moiety. In the benzoic acid dimer these two conformations are equivalent and equi-energetic. In (*S*-1)₄, however, the two conformations are inequivalent and, hence, of different energy. In (*S*-1)₄ each of the four (COOH)₂ moieties can adopt either of these two conformations, which we label a and b, leading to six inequivalent conformations of (*S*-1)₄:

aaaa, aaab, aabb, abab, abbb, bbbb

In aaaa and bbbb, all four (COOH)₂ moieties exhibit equivalent conformations. In aaab and abbb, one (COOH)₂ moiety differs from the other three. In aabb and abab, there are two (COOH)₂ moieties of each conformation; in the former, identical conformations are cis; in the latter they are trans.

We have built and optimized the structures of all six conformations of (*S*-1)₄ at the B3LYP/6-31G* level. Harmonic

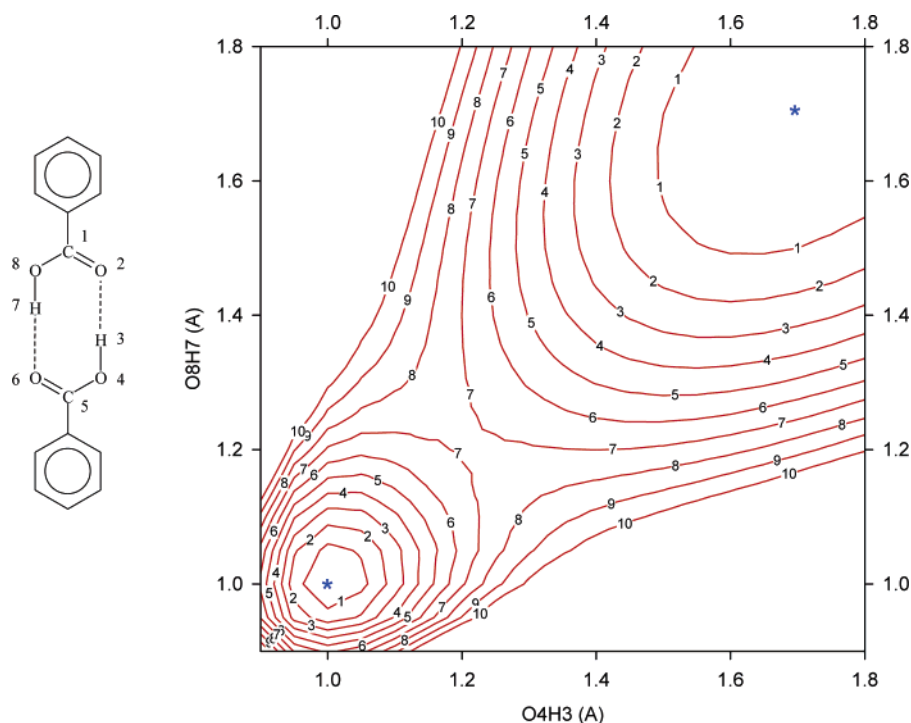


Figure 6. Potential energy surface of $(2)_2$ as a function of the O4H3 and O8H7 distances (Å). The two minima are indicated (*). Contour spacing is 1 kcal/mol.

Table 1. B3LYP/6-31G* Relative Energies, Enthalpies, Entropies, and Free Energies of the Conformations of $(S-1)_4$ ^a

conformer	ΔE	ΔH	ΔS^b	ΔG^c
aaaa (D_4)	0.00	0.00	-8.89 (-2.76)	1.81 (2.62)
aaab (C_2)	1.12	0.84	0.00 (0)	0.00 (0)
aabb (C_1)	1.96	1.88	-6.11 (0)	2.86 (2.86)
abab (D_2)	2.10	1.13	-11.35 (-1.38)	3.67 (4.07)
abbb (C_2)	3.11	2.78	-1.58 (0)	2.41 (2.41)
bbbb (D_4)	3.94	3.82	-12.45 (-2.76)	6.69 (7.50)

^a ΔE , ΔH , and ΔG in kcal/mol; ΔS in cal/mol-K. $T = 298$ K. ^b Residual entropy contributions are in parentheses. ^c Relative free energies including the contributions of residual entropies are in parentheses.

vibrational frequencies have been calculated at the same level in order to confirm that all structures are stable conformations and to enable free energies to be calculated. The relative energies, enthalpies, entropies, and free energies (at 298 K) of the six conformations are listed in Table 1. The two conformations, aaaa and bbbb, in which all $(\text{COOH})_2$ moieties are equivalent and which have D_4 symmetry, differ in energy by 3.94 kcal/mol. In the lower-energy structure, aaaa, every COOH group is nearly coplanar with the adjacent phenyl ring and oriented so that its C=O group is "cis" with respect to the central C-C bond of the biphenyl moiety. In the higher-energy structure, bbbb, all COOH groups are rotated by $\sim 180^\circ$. The four conformations—aaab, aabb, abab, and abbb—in which all $(\text{COOH})_2$ moieties are not equivalent exhibit relative energies intermediate between 0 and 3.94 kcal/mol. The relative energy of aaab, which has C_2 symmetry, is 1.12 kcal/mol. The relative energies of aabb and abab, which have C_1 and D_2 symmetries, respectively, are approximately twice this value. The relative energy of abbb, which has C_2 symmetry, is approximately three times that of aaab. Thus, the relative energy increases quite regularly and linearly with the increasing number of $(\text{COOH})_2$ moieties in conformation b. Surprisingly, the relative free energies are ordered quite differently. The C_2 conformation aaab

has the lowest free energy. The D_4 conformation aaaa is 1.81 kcal/mol higher in free energy. The C_2 conformation abbb is 2.41 kcal/mol above aaab. The C_1 aabb, D_2 abab, and D_4 bbbb conformations lie at increasingly higher free energies. Thus, our calculations predict that the dominant conformation in the room-temperature equilibrium mixture of all six inequivalent conformations is the C_2 conformation aaab, in which three of the four $(\text{COOH})_2$ moieties exhibit the lower-energy conformation a and one exhibits the higher-energy conformation b. (It should be pointed out that the calculations of relative free energies do not include the contributions of the residual entropies⁷ of the six conformations. The aaaa and bbbb conformations are unique structures and have zero residual entropy. However, there are four equivalent structures of the aaab, abbb, and aabb conformations, and each of these conformations has a residual entropy of $R \ln 4$. Likewise, the conformation abab has a residual entropy of $R \ln 2$. These residual entropies and their contributions to the relative free energies are also given in Table 1. Their inclusion increases the free energy difference of the aaaa and aaab conformations to 2.62 kcal/mol, further favoring the aaab conformation in the room-temperature equilibrium mixture of conformations.) The structure of the aaab conformation is shown in Figure 7.

As discussed above, the benzoic acid dimer is planar. In the conformations of $(S-1)_4$, the constraints of forming a cyclic tetramer distort the $C_6-(\text{COOH})_2-C_6$ moieties from planarity. Key structural parameters of the conformations aaaa, aaab, and bbbb of $(S-1)_4$, defining the structures of the $(\text{COOH})_2$ moieties and their orientations with respect to the adjacent phenyl rings, are listed in Table 4 of the SI. The dihedral angles defining the relative orientations of the two phenyl rings of each biphenyl moiety are also given. Overall, the structures of the three

(7) Atkins, P.; de Paula, J. *Physical Chemistry*, 7th ed.; Freeman, W. H.: New York, 2002; pp 672–673.

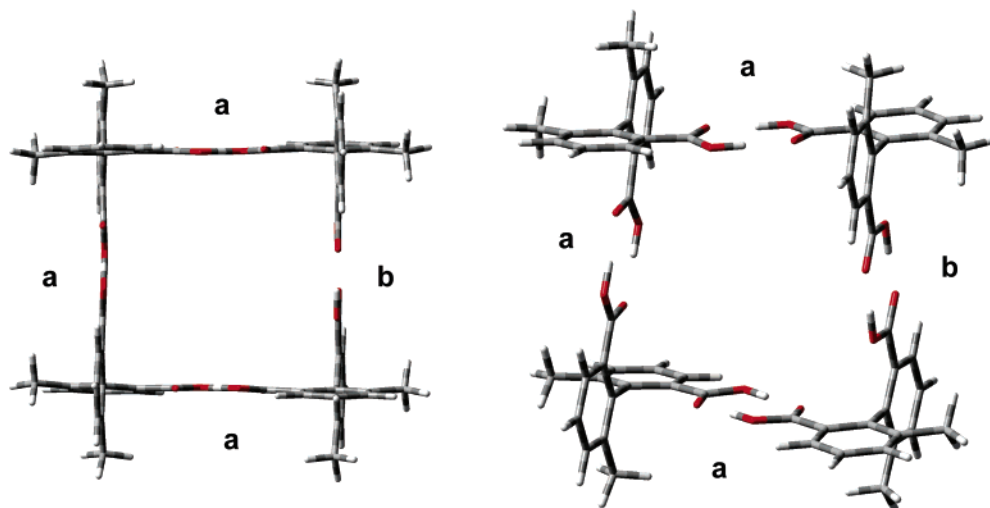


Figure 7. B3LYP/6-31G* structure of the aaab conformer of (*S*-**1**)₄.

conformations of (*S*-**1**)₄ are very similar. In each conformation the two phenyl rings of each biphenyl moiety are approximately perpendicular: the dihedral angles lie within the range 94.6°–92.8°. Each COOH group is planar to a very good approximation and approximately coplanar with the adjacent phenyl ring: dihedral angles defining the COOH–phenyl rotation are <6° from coplanarity. The (COOH)₂ moieties are also approximately planar. The dihedral angles O2O8O6O4, defining the rotation of one COOH group relative to the other, are <4° from coplanarity. Other structural parameters for aaaa, aaab, and bbbb and all structural parameters for aabb, abab, and abbb can be obtained from the Cartesian coordinate geometries given in the SI.

It is of interest to compare our calculated structures to that of (**1**)₄ in crystalline *R*-**1**, obtained by X-ray crystallography.¹ The calculated structure of the aaab conformation of (*S*-**1**)₄ is compared to the structure of the (*R*-**1**)₄ tetramer obtained by X-ray crystallography in Figure 10 of the SI. These structures are substantially different. The X-ray structure is that of an aaaa tetramer (i.e., all (COOH)₂ moieties are in the same tautomeric form). However, it is highly distorted from the *D*₄ symmetry to be expected for an isolated aaaa tetramer molecule. This distortion presumably reflects the presence of strong inter-tetramer, crystal-packing forces (see Figure 4 of ref 1, which shows the packing of tetramers in crystalline (*R*-**1**)₄).

Calculated IR and VCD Spectra of (*S*-1**)₄.** In addition to the B3LYP/6-31G* harmonic vibrational frequencies of the six conformations of (*S*-**1**)₄, we have calculated the harmonic dipole and rotational strengths, with the results tabulated in Table 3 of the SI. Thence, we have predicted the mid-IR, IR and VCD spectra, which can then be compared to the experimental spectra. We begin by focusing on the C=O stretching region of the spectra.

The tetramer (*S*-**1**)₄ contains eight C=O groups and there are therefore eight C=O stretching modes. In the benzoic acid dimer, (**2**)₂, the B3LYP/6-31G* frequencies of the *as* and *s* C=O modes are separated by 45 cm⁻¹. If the coupling between the C=O stretching motions of different (COOH)₂ moieties in (*S*-**1**)₄ is weak compared to the separation of the *s* and *as* modes of each moiety, the modes of the tetramer will consist of two groups of four modes, the higher and lower frequency groups being, to a good approximation, combina-

tions of *as* and *s* (COOH)₂ modes respectively. This is in fact the case, as can be seen from Table 3 of the SI, where the C=O stretching frequencies, dipole strengths, and rotational strengths are listed. In (**2**)₂, the *as* and *s* C=O stretching modes are IR-allowed and -forbidden respectively. A further consequence of the “weak-coupling” regime for C=O stretching modes of (*S*-**1**)₄ is that the dipole strengths of the higher-frequency *as* C=O modes are much larger than those of the lower-frequency *s* C=O modes (Table 2). Thus, only the *as* C=O stretching modes contribute significantly to the IR spectrum. This analysis applies equally to the rotational strengths and the VCD spectrum.

Before proceeding further with the calculated IR and VCD spectra of (*S*-**1**)₄, we return briefly to the experimental IR and VCD spectra of *S*-**1** in CDCl₃. As indicated earlier, the high-concentration IR spectrum exhibits two bands at 1703 and 1683 cm⁻¹ which can be assigned to C=O stretching modes. Support for this assignment is provided by comparison to the C=O stretching frequency of (**2**)₂, 1694 cm⁻¹. The VCD spectrum associated with these bands is bisignate. The frequencies of the two, oppositely signed features are 3 cm⁻¹ above and 9 cm⁻¹ below 1703 cm⁻¹, respectively, crossing zero at 1701 cm⁻¹. At 1683 cm⁻¹, the VCD is weak. This makes clear that the positive and negative VCD bands correspond to two (or more) vibrational transitions, both (all) of which fall within the envelope of the 1703 cm⁻¹ band in the IR spectrum but are not resolved. In order to deconvolve the experimental IR and VCD spectra as far as possible, we have carried out Lorentzian fitting^{5a,b} of the 0.16 M spectra, imposing on our fits the constraint that the same number of bands contribute to the IR and VCD spectra. The results are shown in Figure 8. The frequencies, dipole strengths, rotational strengths, and bandwidths obtained are given in Table 2. Clear evidence of six bands in the range 1710–1670 cm⁻¹ is obtained. Two bands are responsible for the 1703 cm⁻¹ IR feature and the bisignate VCD feature. In addition to the band responsible for the 1683 cm⁻¹ IR feature and the associated weak positive VCD, three extra bands are found at 1709, 1686, and 1672 cm⁻¹. These three bands are not visible in the spectra as peaks or shoulders, but are clearly required to obtain acceptable Lorentzian fits.

In Figures 9 and 10 we plot the dipole strengths, rotational strengths, and frequencies obtained via Lorentzian fitting of the

Table 2. C=O Stretch Frequencies, Dipole Strengths, and Rotational Strengths^a

conformer	ν	D	R
aaaa	1771	4541	4921
	1761	398	-2722
	1761	398	-2722
	1754	0.0	0.0
	1723	0.0	0.0
	1719	4	71
	1719	4	71
	1718	0	0
aaab	1768	3875	2001
	1761	383	-2705
	1756	85	-253
	1751	1917	745
	1722	0	10
	1720	3	17
	1718	2	26
	1705	0	15
aabb	1767	3044	-259
	1758	5	-242
	1754	2089	975
	1749	2004	-601
	1722	3	6
	1719	1	34
	1706	3	86
	1705	0	27
abab	1763	2737	947
	1761	479	-3045
	1752	3935	3657
	1749	77	-1568
	1720	0	0
	1720	1	-35
	1705	0	0
	1704	0	4
abbb	1762	1698	-1205
	1755	1379	1030
	1753	3883	3882
	1748	1086	-3910
	1721	0	-12
	1708	5	106
	1705	1	-15
	1705	2	116
bbbb	1756	0	0
	1753	3801	4092
	1753	3801	4092
	1747	1268	-8744
	1707	0	0
	1705	7	201
	1705	7	201
	1704	0	0
experiment ^b	1709	125 (5.3) ^d	19
	1703 (1705) ^c	1082 (6.5) ^d	953
	1695	398 (7.4) ^d	-1225
	1686	61 (7.4) ^d	-128
	1680 (1682) ^c	652 (8.7) ^d	336
	1672	180 (10.2) ^d	-140

^a ν in cm^{-1} , D in 10^{-40} $\text{esu}^2 \text{cm}^2$, R in 10^{-44} $\text{esu}^2 \text{cm}^2$. ^b Two additional bands at 1744 and 1638 cm^{-1} are also included in the fit to the IR spectrum. ^c VCD fit frequencies that differ from absorption fit frequencies are in parentheses. ^d Line widths (γ values) obtained from the fit to the absorption spectrum are in parentheses.

experimental IR and VCD spectra, together with the calculated values for the six conformations of (*S*-1)₄. In the case of the two D_4 symmetry conformations, aaaa and bbbb, two transitions with large dipole strengths are predicted. For all other, lower-symmetry conformations, four transitions are predicted with significant dipole strengths. The pattern of frequencies and dipole strengths varies with the conformation. Likewise, aaaa and bbbb exhibit two strong oppositely signed rotational

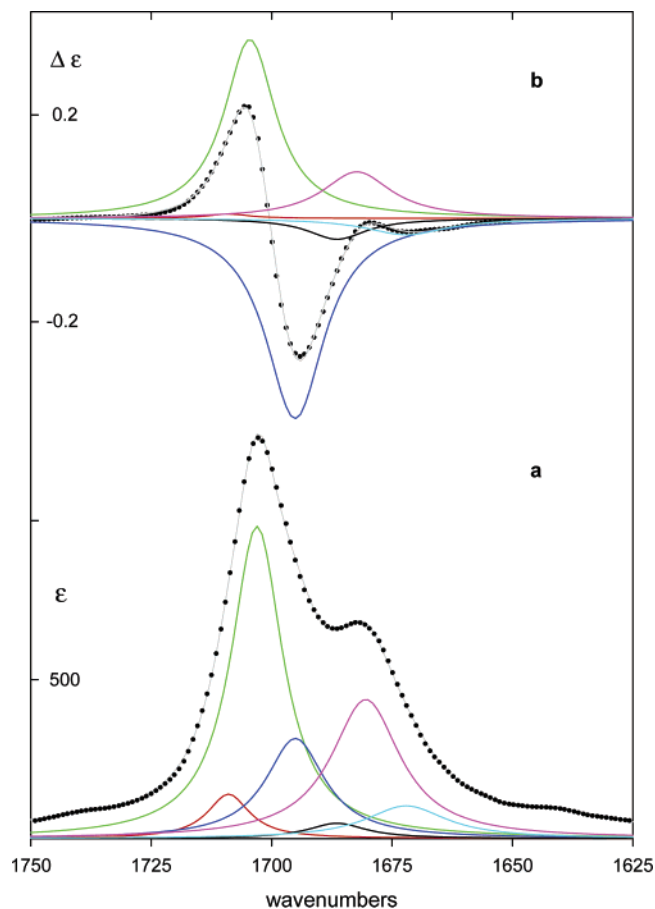


Figure 8. Lorentzian fitting of the experimental CDCl_3 solution IR (a) and VCD (b) spectra of **1** in the C=O stretching region. The experimental spectra ($\bullet\bullet\bullet$) are from Figure 1. The fit is — (gray). The component bands of the fit from 1709 to 1672 cm^{-1} are shown in color.

strengths, while the other conformations exhibit four transitions with significant rotational strengths, the pattern of rotational strengths varying with the conformation. If we exclude the 1709 and 1672 cm^{-1} bands from consideration and focus on the inner group of four experimental bands, comparison to the calculated patterns of frequencies, dipole strengths, and rotational strengths shows that agreement of theory and experiment is exhibited by one, and only one, conformation: the C_2 aaab conformation. Partial agreement is exhibited by the D_4 aaaa conformation: two transitions are predicted to have large dipole strengths and rotational strengths, the latter being opposite in sign; these transitions could be correlated with the experimental bands at 1703 and 1695 cm^{-1} . However, this conformation does not predict significant absorption at lower frequencies and does not account for the 1680 cm^{-1} band in the IR spectrum. For all other conformations—aabb, abab, abbb, and bbbb—the predicted and experimental spectra are in qualitative disagreement.

Our calculations thus support the assignment of four of the experimental IR and VCD bands attributable to C=O stretching modes to the C_2 aaab conformation of (*S*-1)₄. As discussed above, our B3LYP/6-31G* calculations predict that this conformation of (*S*-1)₄ is the lowest in free energy. The comparison of experimental and calculated spectra provides supporting evidence for the reliability of this prediction.

There remain two experimental bands unassigned: at 1709 and 1672 cm^{-1} . Only one conformation, aaaa, exhibits a predicted frequency higher than the highest frequency of aaab.

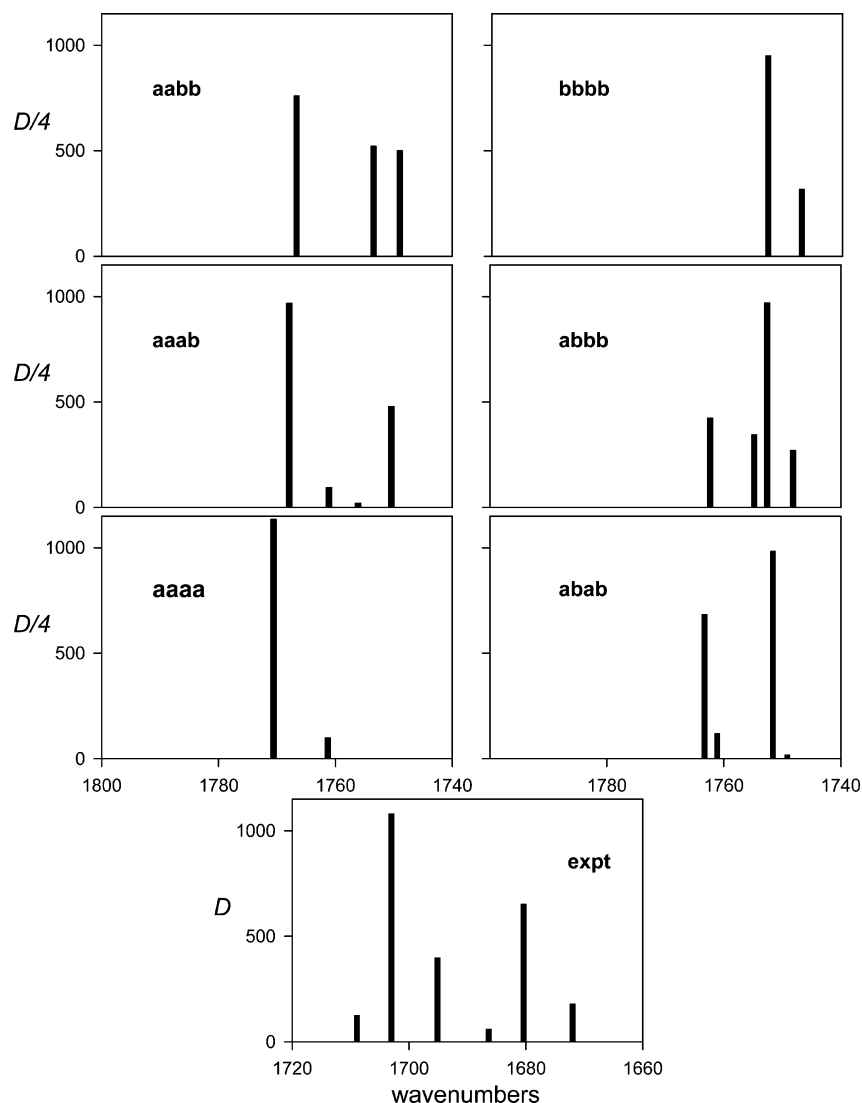


Figure 9. B3LYP/6-31G* frequencies and dipole strengths for the C=O stretching modes of the six conformations of (*S*-**1**)₄ and experimental frequencies and dipole strengths from Lorentzian fitting.

Thus, the only assignment consistent with our calculations for the 1709 cm⁻¹ band would attribute this band to aaaa, in equilibrium with aaab. The positive rotational strength of this band is consistent with this assignment. Likewise, the 1672 cm⁻¹ band could be assigned to the lowest frequency band of one of the conformations aabb, abab, abbb, or bbbb, all of which exhibit negative rotational strengths as observed. However, these assignments would require the relative free energies of the conformations invoked to be significantly lower than predicted, especially in the case of aabb, abab, and bbbb. It seems more likely that the 1709 and 1672 cm⁻¹ bands are overtone or combination bands of aaab and not C=O stretching fundamentals of other conformations.

To further compare theory and experiment, we have synthesized the IR and VCD spectra of the C=O stretching modes of the six conformations of (*S*-**1**)₄ assuming Lorentzian band shapes^{5a,b} whose widths are taken from the Lorentzian fits of the experimental spectra (Table 2). These simulated spectra are shown in Figures 11 and 12, together with the experimental spectra, and further demonstrate the superior agreement with experiment of the predicted spectra for aaab.

We turn now to the mid-IR spectral region to lower frequency of the C=O stretching region. The B3LYP/6-31G* IR and VCD spectra of the aaab conformation of (*S*-**1**)₄ are shown in Figure 13, together with the experimental spectra of the 0.16 M CDCl₃ solution of **1** over the range 1100–1700 cm⁻¹. The density of states is very high in this region: 116 modes are predicted between 1100 and 1700 cm⁻¹. Deconvolution of the experimental spectra, in the same manner as carried out for the C=O stretching region, is therefore impracticable. Calculated and experimental spectra are in qualitative agreement with regard to the principal spectral features, as can be seen in more detail in Figures 2–9 of the SI, where the spectra predicted for the other five conformations of (*S*-**1**)₄ are also shown. The IR spectra in these regions do not vary greatly with conformation. More variation is observed in the VCD spectra, especially in the region 1250–1400 cm⁻¹. The experimental VCD in the 1200–1350 cm⁻¹ range is best reproduced by the calculated VCD for the aaab conformation. Thus, while analysis as detailed as that carried out for the C=O stretching modes cannot be carried out at lower frequencies, comparison of calculated and experimental IR and VCD spectra is at least consistent with the

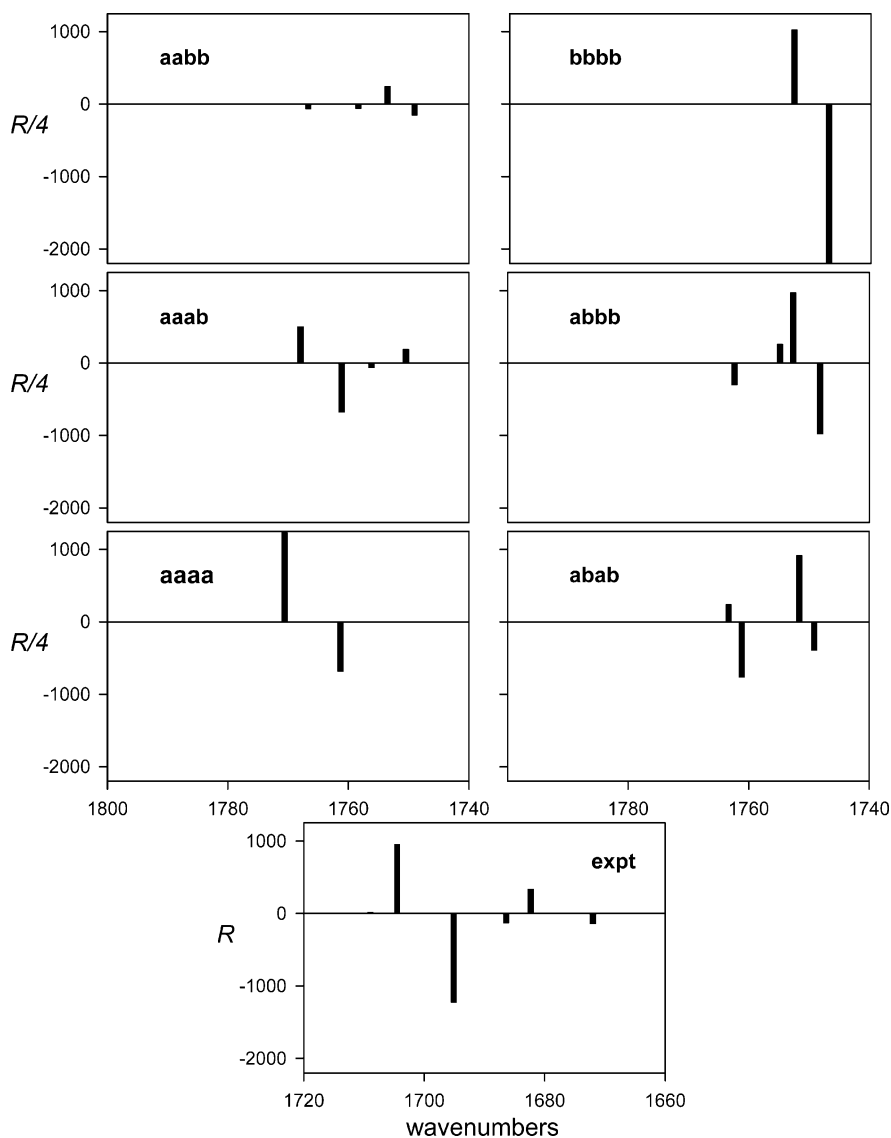


Figure 10. B3LYP/6-31G* frequencies and rotational strengths for the C=O stretching modes of the six conformations of (*S*-1)₄ and experimental frequencies and rotational strengths from Lorentzian fitting.

conclusion that the experimental spectra originate in the *aaab* conformation of (*S*-1)₄.

Discussion

The dimerization of monocarboxylic acids is well-known. The IR spectrum of benzoic acid, **2**, in the nonpolar solvent CHCl₃ over the millimolar-to-molar concentration range reflects this phenomenon. Formation of the (COOH)₂ moiety of the dimer (**2**)₂ is accompanied by lengthening of the C=O and O–H bonds, with the concomitant reduction in C=O and O–H stretching vibrational frequencies. In the case of **2**, the C=O stretch frequency is decreased by 39 cm⁻¹; the O–H stretch frequency is decreased by >300 cm⁻¹, being invisible in the spectrum of (**2**)₂ under the broad, complex C–H stretching absorption. DFT calculations successfully reproduce the IR spectra of **2** and (**2**)₂ and, specifically, the changes in the C=O and O–H stretching modes. It follows that the DFT equilibrium geometries of **2** and (**2**)₂ must successfully reproduce their experimental geometries. C=O and O–H bond lengthenings are predicted to be ~0.02 and ~0.03 Å, respectively. The OH⋯O bridges deviate from linearity by <2°. The H⋯O and

O⋯O H-bonding distances are ~1.67 and ~2.68 Å, respectively.

The IR spectrum of the dicarboxylic acid, *S*-**1**, in CDCl₃ exhibits concentration dependence, demonstrating that aggregation of *S*-**1** takes place with increasing concentration. The changes in the C=O and O–H stretching regions are qualitatively similar to those in solutions of **2**. The C=O stretching band at 1738 cm⁻¹ decreases in intensity with increasing concentration, while bands at 1703 and 1683 cm⁻¹ increase in intensity. The O–H stretching band at 3512 cm⁻¹ decreases in intensity with increasing concentration, without the appearance of any new bands in this region. This latter observation leads to the conclusion that in the oligomer of *S*-**1**, (*S*-**1**)_{*n*}, all carboxyl groups are involved in H-bonded (COOH)₂ moieties (i.e., (*S*-**1**)_{*n*} is a **cyclic** oligomer). Since a cyclic tetramer of *R*-**1**, (*R*-**1**)₄, was observed in crystalline *R*-**1** via X-ray crystallography,¹ it is natural to assign the high-concentration CDCl₃ solution form of *S*-**1** as the cyclic tetramer (*S*-**1**)₄. The same conclusion, that *n* = 4, is independently arrived at by model building. Without considerable distortion of the structure of *S*-**1**, the tetramer

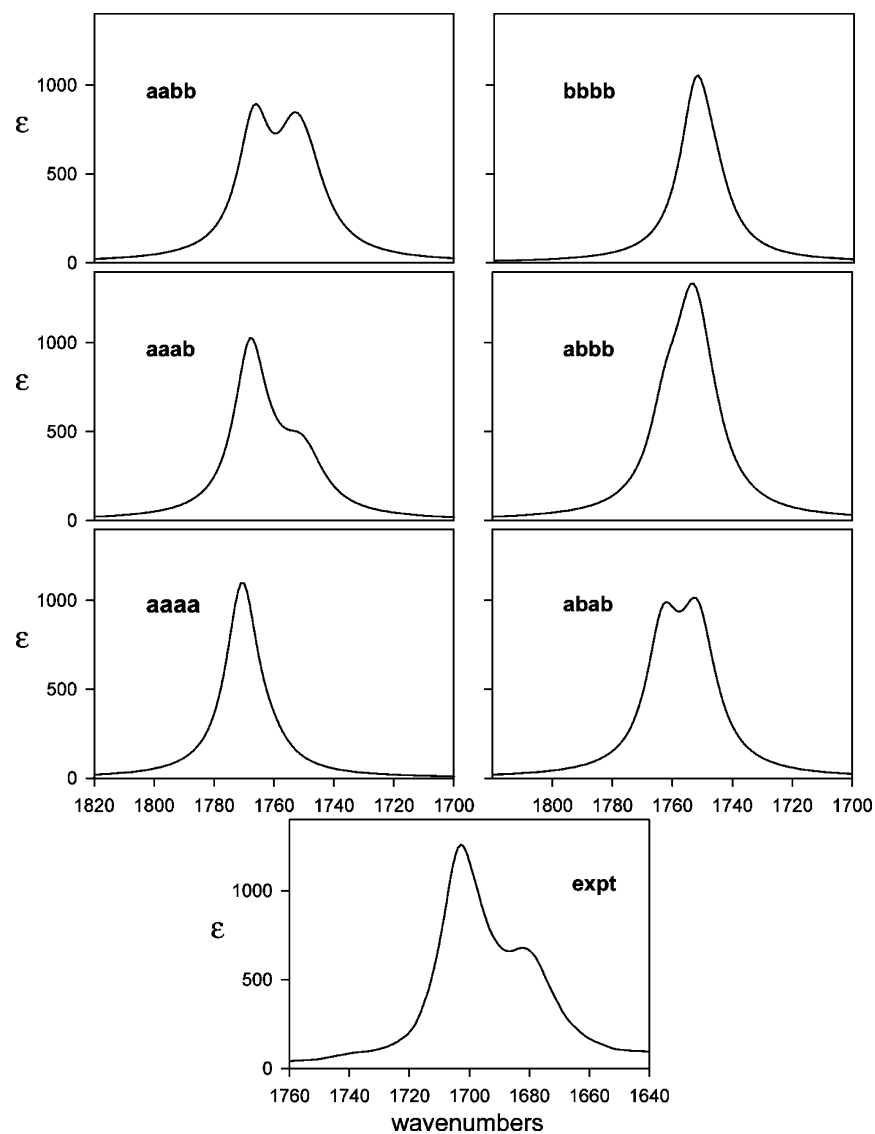


Figure 11. B3LYP/6-31G* IR spectra for the C=O stretching modes of the six conformations of (*S*-**1**)₄ and the experimental IR spectrum (CDCl₃/0.16 M).

constitutes the only way to link the monomers into a cyclic oligomer.

Further evidence for oligomerization of *S*-**1** is provided by the concentration dependence of the VCD spectrum of CDCl₃ solutions of *S*-**1**, especially in the C=O stretching region. As the concentration increases, the C=O stretch VCD changes in shape and increases dramatically in magnitude. The bisignate C=O stretch VCD at high concentrations resembles the “coupled oscillator” VCD of molecules containing two or more C=O groups, first observed in the VCD spectrum of dimethyl tartrate.⁸ The large bisignate C=O stretch VCD in high-concentration CDCl₃ solutions of *S*-**1** clearly originates in interactions between the C=O groups of the oligomeric form of *S*-**1**.

To further define the structure of the cyclic oligomer of *S*-**1** responsible for the high-concentration IR and VCD spectra of *S*-**1** in CDCl₃ solution, we have carried out DFT calculations of the IR and VCD spectra of the cyclic tetramer (*S*-**1**)₄. This requires, initially, the calculation of its equilibrium geometry. Here arises the complication of multiple conformational states

of (*S*-**1**)₄, originating in the two tautomeric conformations of each (COOH)₂ moiety. In the benzoic acid dimer, (**2**)₂, there are two equivalent, equi-energetic tautomeric conformations which can be interconverted either by simultaneous switching of the two O–H H atoms from one side to the other of the (COOH)₂ moiety or by simultaneous rotation of both COOH groups by 180°. The barrier on the former (lower-energy) pathway is 7–8 kcal/mol at the B3LYP/6-31G* level, so that two well-defined conformations are predicted to exist. In (*S*-**1**)₄ each (COOH)₂ moiety must likewise have two conformations. In contrast to (**2**)₂, however, these two conformations are not chemically equivalent and are therefore not equi-energetic. Each of the four (COOH)₂ moieties can be in either of the two tautomeric conformations, resulting in 16 possible structures for (*S*-**1**)₄. Allowing for the symmetry of the system, there are six inequivalent structures. Labeling the two conformations a and b, there are two *D*₄ structures, aaaa and bbbb, one *D*₂ structure, abab, two *C*₂ structures, aaab and abbb, and one *C*₁ structure, aabb. B3LYP/6-31G* calculations for all six structures show that the energy increases approximately linearly with the number of (COOH)₂ moieties in the b conformation. On this basis, one

(8) Keiderling, T. A.; Stephens, P. J. *J. Am. Chem. Soc.* **1977**, *99*, 8061–8062.

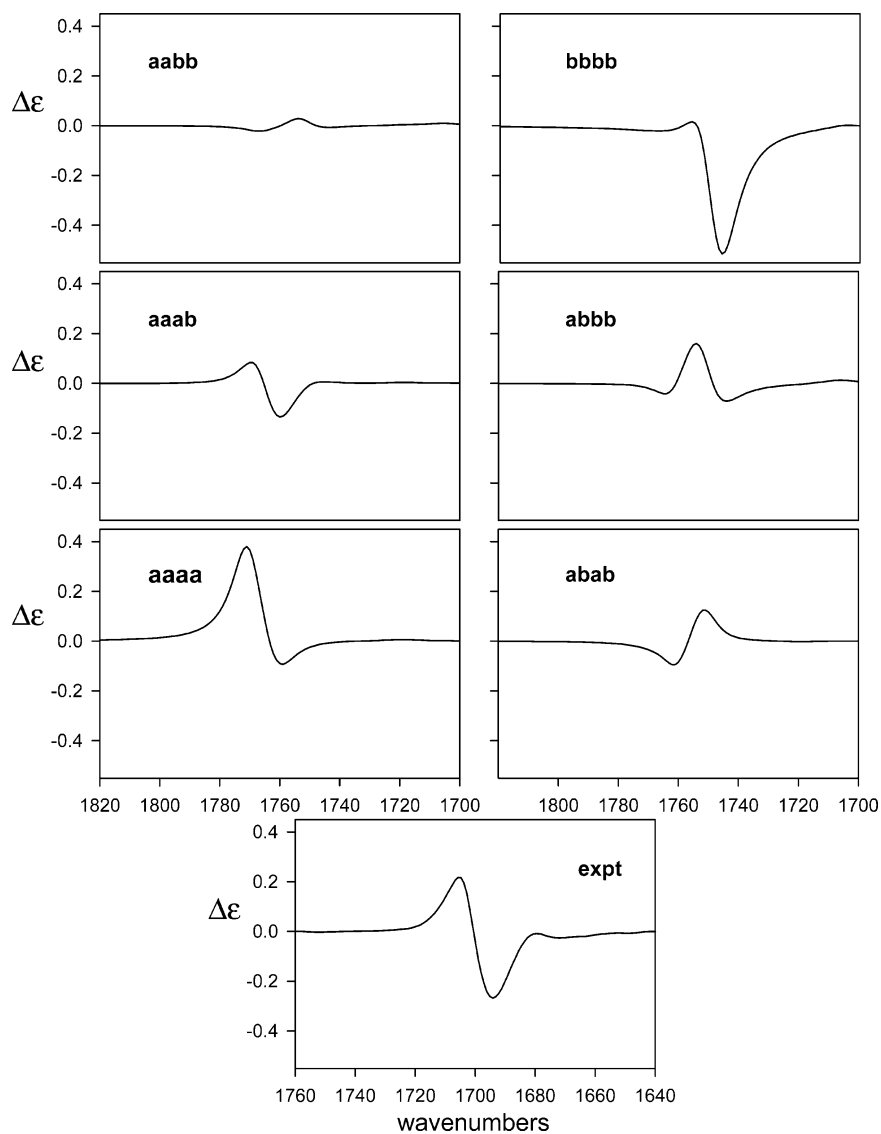


Figure 12. B3LYP/6-31G* VCD spectra for the C=O stretching modes of the six conformations of (*S*-1)₄ and the experimental VCD spectrum (CDCl₃/0.16 M).

might expect the *D*₄ aaaa conformation to predominate. However, on calculation of the relative **free** energies of the six conformations, we find that, at room temperature, the aab conformation has the lowest **free** energy. As a result of the higher entropy of the aaab conformation, its free energy is predicted to be 1.81 kcal/mol (2.62 kcal/mol if residual entropy is included) lower than that of aaaa, despite its 1.12 kcal/mol higher energy. (Similarly, the abbb conformation, whose energy is higher than those of the abab and aabb conformations, has lower free energy than these latter two conformations.)

Comparison of the IR and VCD spectra calculated for the six conformations of (*S*-1)₄ to the high concentration spectra of *S*-1 in CDCl₃ solution permits the identification of the conformation responsible for the experimental spectra. Analysis is simplest for the C=O stretch spectrum. There are eight C=O stretching modes in each conformation. In the “weak-coupling regime”, in which the coupling of the C=O stretches in different (COOH)₂ moieties is small compared to that of the two C=O stretches in the same (COOH)₂ moiety, the C=O stretch modes of (*S*-1)₄ are to a good approximation mixtures of either *s*

(COOH)₂ modes or as (COOH)₂ modes. In a symmetrical (COOH)₂ moiety (such as in (2)₂), the *s* C=O stretch mode has zero dipole moment derivative and is therefore electric dipole-forbidden. It follows that in (*S*-1)₄, in the weak-coupling regime, only four C=O stretch modes will carry significant IR and VCD intensity. This is indeed found to be the case in the calculated IR and VCD spectra of all six conformations of (*S*-1)₄. In the *D*₄ symmetric conformations aaaa and bbbb, even fewer (three) as C=O stretch modes are symmetry-allowed, and two are degenerate. Deconvolution of the C=O stretch IR and VCD spectra identifies six transitions in this region, of which three dominate both IR and VCD. Comparison of the calculated patterns of frequencies, dipole strengths, and rotational strengths to the experimental data shows that qualitative agreement between theory and experiment occurs for one conformation only, aaab, and that it is the inner group of four transitions of the six observed experimentally which are C=O stretch fundamentals. Simulations of the IR and VCD spectra using Lorentzian band shapes and experimental bandwidths confirm this analysis. Thus, we both predict that aaab is the conformation of lowest free energy and find that its predicted C=O stretch

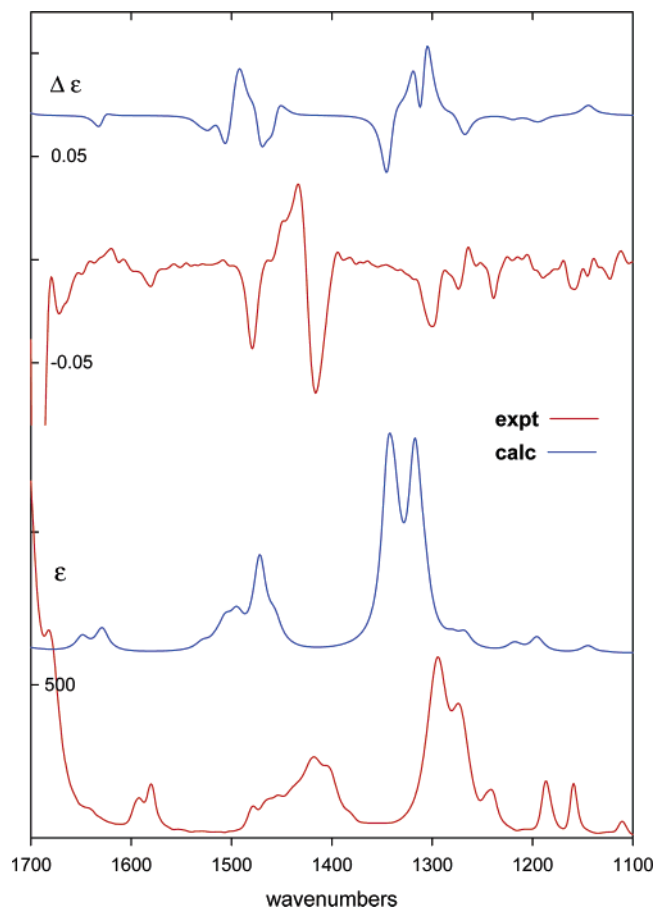


Figure 13. B3LYP/6-31G* IR and VCD spectra of the aaab conformer of (*S*-**1**)₄ and the experimental IR and VCD spectra (CDCl₃/0.16 M) in the region 1100–1700 cm⁻¹.

IR and VCD is in excellent agreement with the experimental spectra, in contrast to the predicted spectra of the five other conformations. Comparison of the IR and VCD spectra at frequencies lower than those of the C=O stretching modes further supports our analysis.

Our analysis of the IR and VCD spectra of (*S*-**1**)₄ is based on B3LYP/6-31G* DFT calculations. In the case of the benzoic acid dimer, (**2**)₂, increasing the number of polarization functions and adding diffuse functions to the basis set led to qualitatively identical and quantitatively very similar IR spectra. We therefore believe that the use of the 6-31G* basis set in the analysis of the IR and VCD spectra of (*S*-**1**)₄ is not a substantial limitation of our calculations. Of course, calculations using larger basis

sets will provide more accurate predictions, and we hope to carry out such calculations at a future time.

IR and VCD spectroscopies are well-suited to the study of the structure of (*S*-**1**)₄ since these techniques are exquisitely sensitive to molecular structure. However, in view of the unexpected conclusion that the dominant conformation of (*S*-**1**)₄ in CDCl₃ solution is the C₂ symmetry conformation aaab, and not the D₄ symmetry conformation aaaa, further studies of this system using other techniques, such as NMR, would be of interest. In addition, further IR and VCD studies of related dicarboxylic acids, such as differently substituted biphenyl dicarboxylic acids and binaphthyl dicarboxylic acids, which are likely to exhibit analogous behavior would also be of interest. Such studies are planned.

Our analysis of the structure of (*S*-**1**)₄ is the first to use VCD spectroscopy, coupled with DFT calculations, to elucidate the structure of a supramolecular species. Our work demonstrates the potential utility of VCD spectroscopy in the field of supramolecular chemistry.²

Acknowledgment. We are grateful to Dr. Tomáš Kotrba for assistance in the early stages of this work. We are extremely grateful to Angela Loh and Richard Readings of Hewlett-Packard Corp. for time and assistance on HP computing systems, without which this work would not have been possible, and to Dr. James R. Cheeseman of Gaussian Inc. for porting Gaussian 03 to these computers. The work was supported by the research grants 1K03012 and MSM6046137307 to M.U. from the Ministry of Education, Youths and Sports of the Czech Republic, and by grant CHE-0209957 to P.J.S. from the U.S. National Science Foundation. We also acknowledge the support of the USC Center for High Performance Computing and Communications, HPCC.

Supporting Information Available: Harmonic frequencies, dipole strengths, and rotational strengths for **2**, (**2**)₂, and the six conformations of (*S*-**1**)₄. Structural parameters of **2**, (**2**)₂, and (*S*-**1**)₄. Comparisons of B3LYP/6-31G* calculated IR and VCD spectra and experimental IR and VCD spectra of (*S*-**1**)₄. Comparisons of B3LYP/6-31G* calculated dipole and rotational strengths and experimental IR and VCD spectra of (*S*-**1**)₄. Comparison of the calculated structure of (*S*-**1**)₄ and the X-ray structure of (*R*-**1**)₄. Cartesian coordinates of optimized geometries. Methods section. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA050483C