Molecular Bilateral Symmetry of Natural Products: Prediction of Selectivity of Dimeric Molecules by Density Functional Theory and Semiempirical Calculations

Tatyana Voloshchuk, Nicola S. Farina, Orrette R. Wauchope, Magdalena Kiprowska, Paul Haberfield, and Alexander Greer*

Department of Chemistry, Graduate School and University Center and The City University of New York (CUNY), Brooklyn College, Brooklyn, New York 11210

Received March 10, 2004

A literature survey and theoretical calculations have been applied to explore bilateral symmetry in natural product systems. Molecular bilateral symmetry is defined to include C_2 (sigma plane or axis), C_s, and C_{2v} point groups in molecules. Natural products that possess chirality in the form of C_2-axes or sigma planes of symmetry are present in higher proportions (69%) compared to molecules bearing achiral C_s or C_{2v} point groups (14% and 16%, respectively). Density functional theoretical and semiempirical calculations indicate that the dimers 3,3′-dibromo-5,5′-[*N*-(2-(3-bromo-4-hydroxyphenyl)ethyl)-2-hydroxyiminoacetamide]biphenyl-2,2′-diol (**1**), (*S*,*S*)-1,2-bis(2-amino-3*H*-imidazol-4-yl)-(*R*,*R*)-3,4-bis(1*H*-pyrrole-2-amido)-cyclobutane (**2**), 2-oxo-dimethyl-1,3-bis(3,4-dibromobenzene-1,2-diol) (**11**), 1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,6-diene-3,5-dione (**12**), and bis(5-isopropyl-8-methylazulene)methane (**13**) evolve more energy per connecting bond than the corresponding trimers or tetramers would. This we propose is a guiding parameter that may adjust molecule growth. The corresponding trimers, tetramers, or higher oligomers of **1**, **2**, and **11**−**13** appear to represent "missing" compounds in nature. Natural products **1**, **2**, and **11**−**13**, having 3-fold and higher levels of symmetry, would founder on the lack of a facile method of synthesis and on the prohibitively high-energy costs caused by steric crowding at their core.

The number of natural product molecules that possess bilateral symmetry seems to exceed what one would expect on the basis of pure chance. This observation is reminiscent of the fact that many animal phyla exhibit apparent bilateral body symmetry, wherein a plane of symmetry generates two mirror image halves. *Molecular bilateral symmetry* may be defined as including C_2 (sigma plane or axis), C_s, and C_{2v} point groups in molecules (Figure 1). In addition, a molecule, such as an optically active biphenyl, can adopt two mirror image halves in a transition state to racemization.[1]

Representative examples of naturally occurring molecules that possess C_2 symmetry include **1**−**5**. Compounds **6**−**10** represent examples of natural products with C_s symmetry, and compounds **11**−**15** represent those with C_{2v} symmetry. Natural products with biaryl or macrocyclic groups are common. It should be noted that some natural products can be regarded as dimers, but do not possess C_2, C_s, or C_{2v} symmetry. Natural products that fall into this class, e.g., **16**−**22**, were not considered in further detail since symmetry elements may resemble bilateral symmetry, but are often not obvious by looking at the compound. The biosynthesis of epoxyquinol **21**[2] and the laboratory synthesis of usnic acid **22**[3] serve as examples, which constitute a lower symmetry for the dimer.

The literature appears to be devoid of studies that address a possible role for the bilateral symmetry of natural products or a discussion that compares the physical and chemical properties of such molecules with those that lack this symmetry. Little is known on why bilaterally symmetrical compounds come into being or why nature uses dimeric molecules. Bilateral structural features have not been correlated with existing biological data, nor are indices available to classify these compounds. It might be useful to discover a connection between this symmetry and some function or property of these molecules. In a similar vein, the practice of organic chemists doing synthesis often

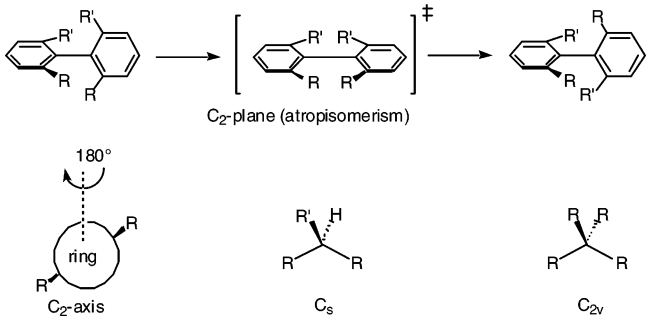


Figure 1. Examples of molecular bilateral symmetry.

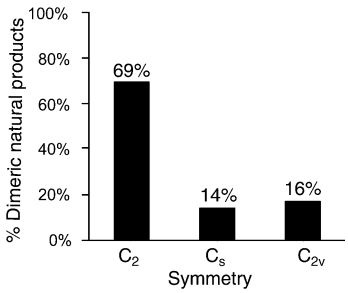
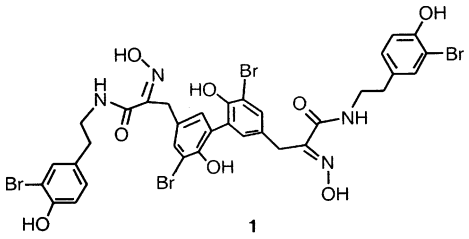


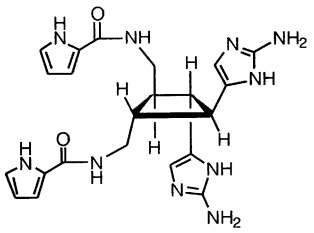
Figure 2. Correlation of point groups of bilaterally symmetrical natural products. In this survey, a dimer is defined as a molecule having central bond(s) that join two segments of atoms of the same kind or where monomers are positioned between a branching point. Atropisomeric natural products feature C_2-axes or sigma planes of symmetry where compounds expressed chirality or were isolated as racemic mixtures. One (C_s) or two planes (C_{2v}) of symmetry lead to an absence of chirality as does a rotoreflection axis S_2. Natural products that possess C_{2h} or S_2 or higher symmetry appear to be rare (<1%). Dimers of D_n, $C_{\infty h}$, D_{nd}, D_{nh}, T_d, O_h, and I_h symmetry were not found in the literature search.

* To whom correspondence should be addressed. Tel: 718-951-5475. Fax: 718-951-4607. E-mail: agreer@brooklyn.cuny.edu.

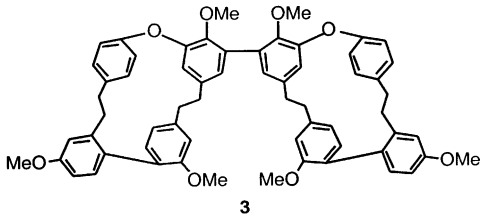
10.1021/np049899e CCC: $27.50
Published on Web 07/08/2004



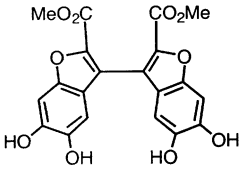
1



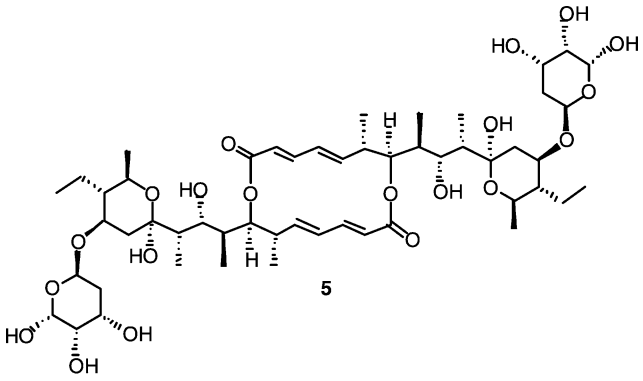
2



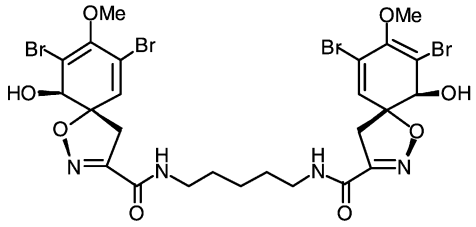
3



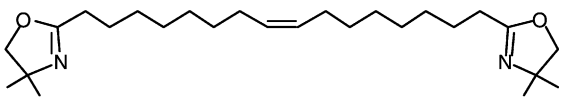
4



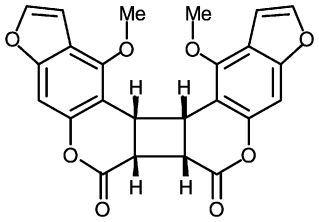
5



6



7



8

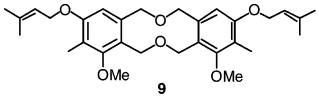
uses "mirror" image symmetry to simplify molecule construction.[4] For instance, connecting two halves prepared from common monomers simplified the synthesis of squalene[5] and isochrysohermidin.[6] The term for mirror image symmetry is used here as an example of a molecule having a σ plane (S_1 axis of symmetry) or of a molecule having two chiral centers (an example of *dl*/meso) considering a planar "biphenyl" system.

A literature survey and theoretical calculations have been applied to explore bilateral symmetry in natural product systems. In this work we propose to ask several questions that may lead to a better understanding of the extent of bilateral symmetry among natural products. Is there a correlation between different symmetries and the source of the molecule in the plant and animal kingdom? Are there energy costs and/or advantages associated with the assembly of molecules of this kind? Can a literature survey and computational methods shed light on these possibilities?

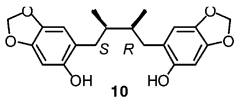
Results and Discussion

Bilaterally Symmetrical Chemicals. We sought a basis to tabulate natural product bioactivities and the source from which the substances were isolated. Some 3000 articles were examined in which approximately 7% included molecules possessing bilateral symmetry.[7] A trend emerged when comparing the symmetry elements of the dimeric natural products comprised of the 7% that possess bilateral symmetry (Figure 2). Molecules that potentially possess chirality in the form of C_2-axes or sigma planes of symmetry are present in high proportions (69%) compared to molecules bearing achiral C_s or C_{2v} point groups (14% and 16%, respectively). Natural products with biaryl or macrocyclic groups are common. The data can be organized into four groups based on the origin of these compounds: plant (65%), animal (22%), bacterial (8%), and fungal (4%). This does not appear to represent a preference for bilateral symmetry in plants, but rather a prevalence of natural products reported in the literature from plant species. To our knowledge, no correlation has been published between biochemical assembly and molecules that exhibit bilateral symmetry.

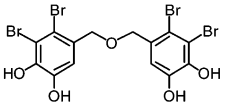
Bond Energy Data. Turning to the subject of the assembly of bilaterally symmetrical molecules, we have compared the heats of formation of dimers, trimers, and tetramers. To evaluate bond energy costs, a test set of



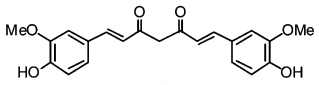
9



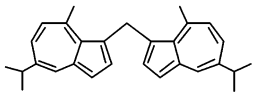
10



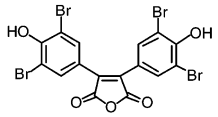
11



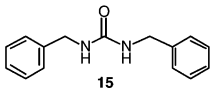
12



13



14



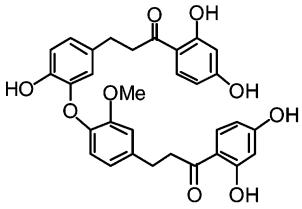
15

hydrocarbons, **23**−**28**, was explored as a first step. The B3LYP/D95**, AM1, and ONIOM(B3LYP/D95**:AM1) methods employed here compare well with literature C−C bond distances and bond energies (Table 1).[8]

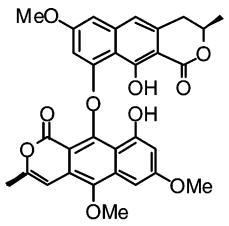
Table 1. Calculated and Experimental Values of Hydrocarbon Structural Parameters and Bond Energies

compound	C−C (lit.)[a]	C−C (calcd)	bond energy (lit.)[b]	bond energy (calcd)[c]
CH_2CH-CN **23**	1.438	1.439[d]	132.1	135.3[e]
		1.419[f]		127.8[g]
Ph−Ph **24**	1.48	1.488[d]	118.0	114.8[e]
				115.0[h]
				98.7[g]
Ph_3C-CPh_3 **25**	1.72	1.612[f]	16.6	42.9[g]
				19.3[i]
H_3C-CH_3 **26**	1.535	1.538[d]	89.7	87.0[e]
				95.8[h]
		1.500[f]		77.3[g]
				90.0[i]
$C_2H_5-C_2H_5$ **27**	1.531	1.514[f]	87.2	62.1[g]
				85.8[i]
				66.1[j]
		1.537[k]		85.7[l]

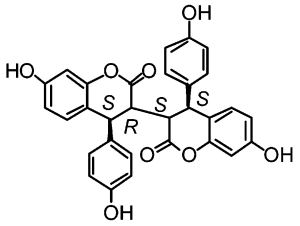
[a] Literature values of bond lengths (distances in Å).[8] [b] Literature values of bond energies (kcal/mol at 298 K).[8] [c] C−C bond energies calculated from a homolytic reaction of unpairing of electrons. [d] B3LYP/D95** optimized geometry. [e] Unrestricted B3LYP/D95** calculation. [f] AM1 optimized geometry. [g] Unrestricted AM1 calculation. [h] Restricted B3LYP/D95** calculation of the fragments optimized with the resulting species isolated at a fixed distance of the literature C−C bond distance value + 1.0 Å. [i] Unrestricted B3LYP/D95**//AM1 calculation. [j] Restricted HF/3-21G//AM1 calculation. [k] ONIOM(B3LYP/D95**:AM1) optimized geometry. The methylene groups were computed with the DFT part and the remaining methyl groups with the semiempirical part. [l] Restricted ONIOM(B3LYP/D95**:AM1) calculation optimized with the resulting species isolated at a fixed distance of the literature C−C bond distance value + 1.0 Å.



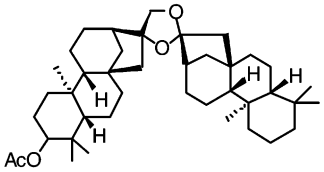
16



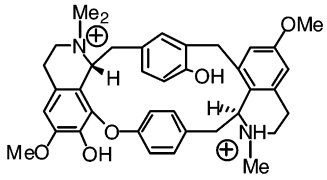
17



18

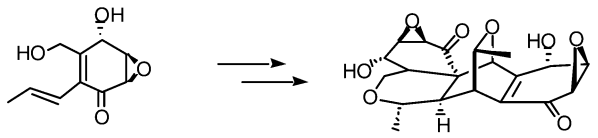


19

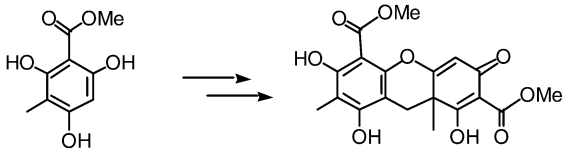


20

Experimental C−C single bond distances range from 1.44 to 1.64 Å, and the longest C−C bond known is 1.72 Å (Table 1). B3LYP/D95** and ONIOM(B3LYP/D95**:AM1) geometry optimizations provide good agreement with a 0.008 Å discrepancy. The AM1 method provides errors that range from 0.017 to 0.108 Å. A stabilizing effect is computed with unrestricted B3LYP/D95** calculations in the carbon−carbon single bond of $H_2C{=}CH-CN$ (**23**), which is largely due to an inductive interaction between the triple bond and the cyano group. The experimental energy for the homolysis of the C−C single bond in biphenyl (**24**) correlates with that found from unrestricted B3LYP/D95** calculations for the reaction Ph−Ph → 2 Ph and from restricted B3LYP/D95** calculations where the fragments were optimized with the resulting species isolated at a fixed distance of



21



22

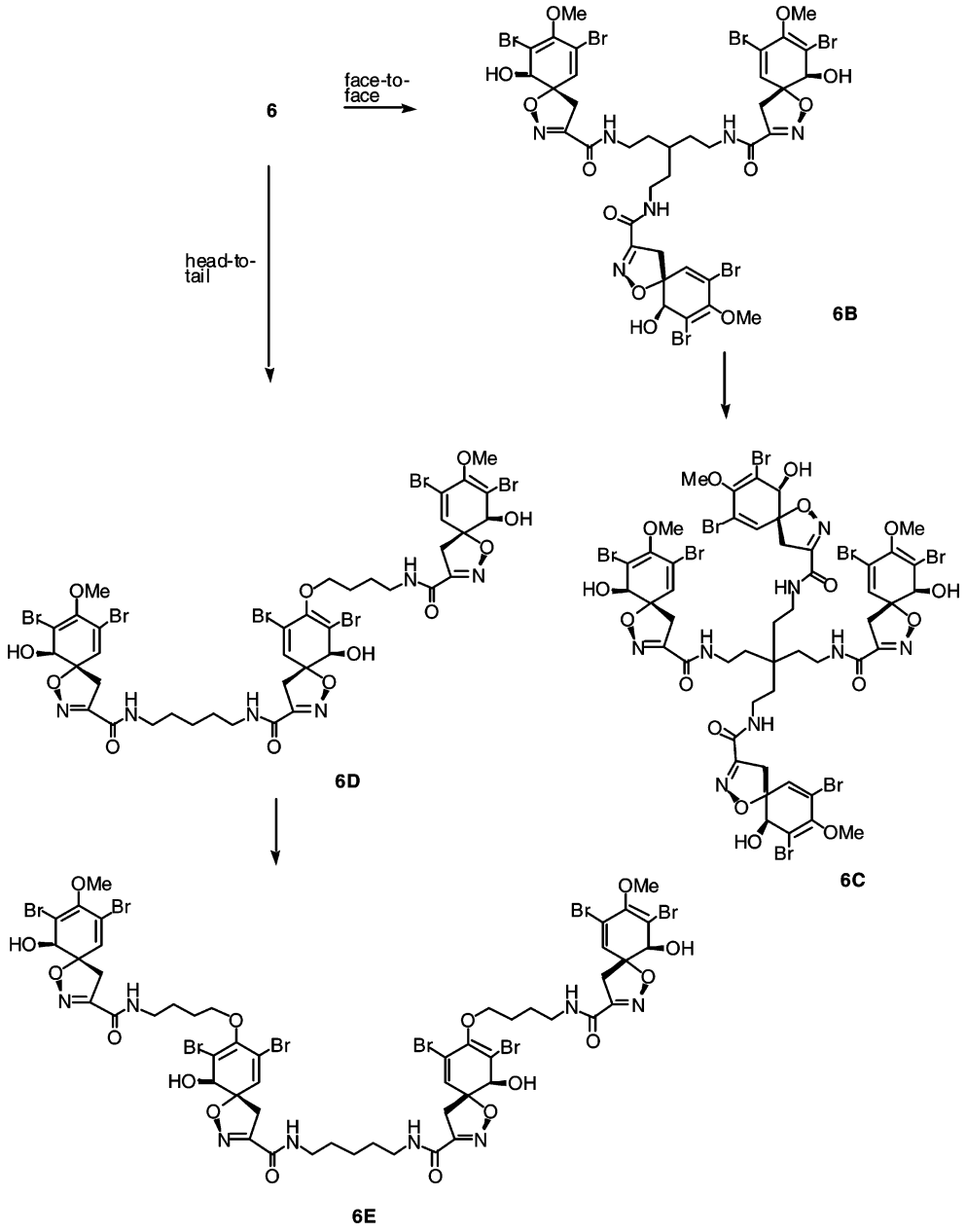
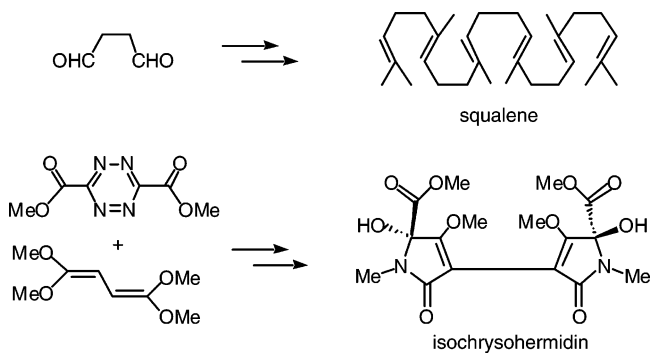


Figure 3. Example of face-to-face contact between monomer units and an alternative head-to-tail connectivity.

the literature C−C bond distance + 1.0 Å. AM1 calculations underestimate the bond energy of the central C−C bond in biphenyl by 19.3 kcal/mol. The bond distances and bond energies can be used to gauge steric congestion in molecules. A classic example of steric crowding in hexaphenylethane, Ph_3C-CPh_3 (**25**), is due to the close proximity of the phenyl groups to one another.[9] This leads to a C−C bond energy of 16 kcal/mol so that homolysis to yield triphenylmethyl radicals and 1-diphenylmethylene-4-trityl-2,5-cyclohexandiene becomes possible.[10] The C−C bond energy of hexaphenylethane **25** is modeled well with the unrestricted B3LYP/D95**//AM1 method where steric interactions, enhanced C−C bond distances, and reduced bond energies are evident when compared to the linear alkanes Me−Me, **26**, and Et−Et, **27**. The degree of branching and distribution of adjacent groups, however, can lead to a reduced C−C bond stability, with steric repulsion giving rise to perturbed systems.[11,12]

Relief of Steric Crowding to Derive a Dimer-Based Mechanism. To explore bilateral symmetry as a possible structure concept, the research focused on predicting the bonding properties of the monomer building blocks in a series of natural product dimers: 3,3′-dibromo-5,5′-[*N*-(2-(3-bromo-4-hydroxyphenyl)ethyl)-2-hydroxyiminoacetamide]-biphenyl-2,2′-diol (**1**),[13] (*S*,*S*)-1,2-bis(2-amino-3*H*-imidazol-4-yl)-(*R*,*R*)-3,4-bis(1*H*-pyrrole-2-amido)cyclobutane (**2**),[14] 2-oxo-dimethyl-1,3-bis(3,4-dibromobenzene-1,2-diol) (**11**),[15]



1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,6-diene-3,5-dione (**12**),[16] and bis(5-isopropyl-8-methylazulene)methane (**13**).[17]

The carbon numbering scheme in which the bond distances and bond energies are described is shown in Table S1 (Supporting Information). AM1, B3LYP/D95**//AM1, HF/3-21G//AM1, and ONIOM calculations predicted that C−C bond strengths decreased incrementally, suggesting a progression for the facile homolysis in the order tetramer > trimer > dimer.

The computed data provide evidence that bond energies of dimeric natural products are enhanced compared to the hypothetical branched trimers and tetramers. Natural products that possess bilateral symmetry share in common an apparent face-to-face contact between the two like monomer units. Figure 3 illustrates an alternative head-to-tail connectivity, which is conceivable[18] even though there exists a tendency for dimer natural product biosynthesis to terminate with a face-to-face contact. For molecules to achieve 3-fold or higher symmetry, and to consider an associated structure−function relationship, branching at the core of a dimer to yield trimers and tetramers may relate symmetry elements to shed light on the mechanism of natural product stability (Table S1; Supporting Information). Compounds **6**, **6B**, and **6C** illustrate such a bonding motif with successive increases in the degree of symmetry. Head-to-tail binding of natural product monomers to give hypothetical trimers (e.g., **6D**) and tetramers (e.g., **6E**) yields molecules that would bear partial or full bilateral symmetry as displayed in the building blocks for common biopolymers from amino acids, nucleic acids, and glucose. For example, glutathione involves a head-to-tail connectivity of three amino acids where the corresponding amide bond dissociation energies do not decrease in this case. The head-to-tail connection pathway may be competitive, but is not common among bilaterally symmetrical natural products.

Conclusion

Structure elucidation is often sought for natural compounds from extracts that exhibit biological activity. The corresponding trimers, tetramers, or high oligomers often represent "missing" compounds from these extracts. Approximately 7% of natural products possess bilateral symmetry; however, the complexity of the molecules and their abundance across a range of structures have hampered the emergence of an understanding that connects phenomena that may be related. Whether bilateral symmetry is thought of as a structure or a property concept, the presence of this symmetry appears to apply to a manifold of natural products and the building blocks from where they arise. For instance, the coupling of aryl rings to give biaryl provides an example of this construction.[19]

The computed data provide evidence that bond energies of dimeric natural products (**1**, **2**, and **11**−**13**) are enhanced compared to the hypothetical branched trimers and tetramers. The finding may be intuitively obvious, but allows conjecture to be made about the biogenesis of the compounds and accounts for the apparent absence of natural products possessing higher symmetry. Adding a further degree of symmetry would contribute successively smaller increments of advantage. Molecules having such 3-fold and higher levels of symmetry would founder on the lack of a facile method of synthesis and on the prohibitively high-energy costs caused by steric crowding at their core. Similar through-space interactions involving steric strain have been recognized in synthetic polymer systems where destabilization is observed when branched carbon−carbon bonds connect monomer units.

Experimental Section

Literature Survey. A survey of natural product structures was derived from the SciFinder Scholar database (Chemical Abstracts Service, Columbus, OH) and by analysis of the *Journal of Natural Products* from 1996 to present. The survey emphasized tabulating point groups and symmetry elements among natural products that may or may not be related compounds. Standard procedures were used in the analysis of natural product symmetries.[20] Natural products were tabulated according to their potential adaptation into the C_{2v}, C_2, and C_s symmetries.

Computations. Density functional theory (DFT), Hartree−Fock (HF), and semiempirical calculations were performed to model the structural features and bond energies.[21,22] Geometries were optimized with the B3LYP/D95**, B3LYP/6-31G*,[23−25] and AM1 methods,[26] or the ONIOM method[27,28] using a two-layer DFT (B3LYP/D95**) and semiempirical (AM1) approach. The conformational preference of molecules was investigated with the empirical MM2 method.[29] For the ONIOM calculations the core of each molecule with surrounding H atoms was calculated at the B3LYP/D95** level. The remaining atoms within each molecule for the ONIOM calculations were calculated at the AM1 level. Since calculations on some large natural product molecules are not yet practical with ab initio and density functional theory, our goal with the ONIOM method was to model C−C bond energies and intramolecular interactions of the natural products. The two-layer ONIOM method reproduced experimental energies of bimolecular reaction systems with good accuracy and compares well with benchmark ab initio calculations.[28] The DFT, AM1, and ONIOM methods yielded bond energies and geometries that correlated well with experimental values and with the relative stability of hydrocarbons.[30] A possible source of error in the calculations is the conformational choice of the ground-state geometry compared to the separated "homolyzed" units. The computational methods used here for bond separation energies via fixed distance calculations can provide a cancellation of error for bond energies in complex natural molecules. This study examined bond distances and bond energies of hydrocarbons and natural product dimers and oligomers in the gas phase.

Supporting Information Available: Descriptions of the computations for the bonding properties of the monomer building blocks in a series of dimers, trimers, and tetramers of **1**, **2**, and **11**−**13**. The Supporting Information is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Eliel, E. R.; Wilen, S. H.; Doyle, M. P. *Basic Organic Stereochemistry*; Wiley-Interscience: New York, 2001; pp 608−642.
(2) Kakeya, H.; Onose, R.; Koshino, H.; Yoshida, A.; Kobayashi, K.; Kageyama, S.-I.; Osada, H. *J. Am. Chem. Soc.* **2002**, *124*, 3496−3497.
(3) Barton, D. H. R.; Deflorin, A. M.; Edwards, O. E. *J. Chem. Soc.* **1956**, 530−534.
(4) Boger, D. L. *Modern Organic Synthesis*; TSRI Press: La Jolla, CA, 1999; pp 431−438.
(5) Johnson, W. S.; Wethermann, L.; Bartlett, W. R.; Brocksom, T. J.; Li, T.-T.; Faulkner, D. J.; Petersen, M. R. *J. Am. Chem. Soc.* **1970**, *92*, 741−742.
(6) Boger, D. L.; Baldino, C. M. *J. Am. Chem. Soc.* **1993**, *115*, 11418−11425.
(7) An additional ∼10% of the 3000 articles included molecules that may be regarded as "derivatized dimers", but do not possess bilateral symmetry. It is amusing to speculate that this ∼10% case once represented bilaterally symmetrical molecules that subsequently evolved to produce more specialized activity by losing exact bilateral symmetry.
(8) Zavitsas, A. A. *J. Phys. Chem. A* **2003**, *107*, 897−898.
(9) Skinner, K. J.; Hochster, H. S.; McBride, J. M. *J. Am. Chem. Soc.* **1974**, *96*, 4301−4306.
(10) Lankamp, H.; Nauta, W. T.; MacLean, C. *Tetrahedron Lett.* **1968**, 249−254.
(11) Sayar, M.; Stupp, S. I. *Macromolecules* **2001**, *34*, 7135−7139.
(12) Shchipunov, Y. A.; Hoffmann, H. *Langmuir* **1998**, *14*, 6350−6360.
(13) Franklin, M. A.; Penn, S. G.; Lebrilla, C. B.; Lam, T. H.; Pessah, I. N.; Molinski, T. F. *J. Nat. Prod.* **1996**, *59*, 1121−1127.

(14) Shen, X.; Perry, T. L.; Dunbar, C. D.; Kelly-Borges, M.; Hamann, M. T. *J. Nat. Prod.* **1998**, *61*, 1302−1303.
(15) Kurihara, H.; Mitani, T.; Kawabata, J.; Takahashi, K. *J. Nat. Prod.* **1999**, *62*, 882−884.
(16) Masuda, T.; Matsumura, H.; Oyama, Y.; Takeda, Y.; Jitoe, A.; Kida, A.; Hidaka, K. *J. Nat. Prod.* **1998**, *61*, 609−613.
(17) Seo, Y.; Rho, J.-R.; Geum, N.; Yoon, J. B.; Shin, J. *J. Nat. Prod.* **1996**, *59*, 985−986.
(18) Dasent, W. E. *Nonexistent Compounds*; Marcel Dekker: New York, 1965.
(19) A reviewer posed a question on why dimeric molecules may be preferred over their monomeric compounds in the biological system. A reaction might proceed in the sequence monomer → dimer, since the economy for doubling molecular weight and complexity comes at a lower cost in energy than would the association of small fragments in a manner analogous to using bricks and mortar.
(20) Eliel, E. R.; Wilen, S. H.; Doyle, M. P. *Basic Organic Stereochemistry*; Wiley-Interscience: New York, 2001; pp 45−64.
(21) Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: New York, 1999; pp 177−193.
(22) Stewart, J. J. P. *Semiempirical Molcular Orbital Methods*; VCH Publishers: New York, 1990; Vol. 1.
(23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian: Pittsburgh, PA, 1998.
(24) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(25) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(26) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(27) Matsubara, T.; Maseras, F.; Koga, N.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 2573−2580.
(28) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357−19363.
(29) Allinger, N. L.; Yuh, Y. H.; Li, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8582.
(30) Repasky, M. P.; Jayaraman, C.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601−1622.

NP049899E