185

This is a chapter from

Concepts in Chemistry:
A Contemporary Challenge

Edited by
Dennis H. Rouvray
University of Georgia, USA

RESEARCH STUDIES PRESS LTD.
Taunton, Somerset, England

JOHN WILEY & SONS INC.
New York · Chichester · Toronto · Brisbane · Singapore

1997

RESEARCH STUDIES PRESS LTD.
24 Belvedere Road, Taunton, Somerset, England TA1 1HD

Marketing and Distribution:

Australia and New Zealand:
Jacaranda Wiley Ltd.
GPO Box 859, Brisbane, Queensland 4001, Australia

Canada:
JOHN WILEY & SONS CANADA LIMITED
22 Worcester Road, Rexdale, Ontario, Canada

Europe, Africa, Middle East and Japan:
JOHN WILEY & SONS LIMITED
Baffins Lane, Chichester, West Sussex, UK, PO19 1UD

North and South America:
JOHN WILEY & SONS INC.
605 Third Avenue, New York, NY 10158, USA

South East Asia:
JOHN WILEY & SONS (SEA) PTE LTD.
37 Jalan Pemimpin 05-04
Block B Union Industrial Building, Singapore 2057

Library of Congress Cataloging-in-Publication Data

Available

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN 0 86380 200 1 (Research Studies Press Ltd.)
Identifies the book for orders except in America.

ISBN 0 471 96555 3 (John Wiley & Sons Inc.)
Identifies the book for orders in USA.

Printed in Great Britain by SRP Ltd., Exeter

CHAPTER 9

The Measurement of Symmetry and Chirality: Conceptual Aspects

D. Avnir, O. Katzenelson, S. Keinan, M. Pinsky, Y. Pinto, Y. Salomon *and* H. Zabrodsky Hel-Or

1. The Primary Concept: Symmetry as a Continuous Structural Property

1.1 Thesis and Motivation

The concept of symmetry has attracted virtually all domains of intellectual activity and has strongly influenced the sciences and the arts.[1-3] It has functioned as a condensed language for the description and classification of shapes and structures; as an identifier of inherent correlations between structure and physical properties of matter; and as a guideline in artistic and practical aesthetic design. This study of symmetry is based on the following thesis:

Thesis: The 'Symmetry of Nature', from molecular scales and up, is a rarely attainable idealized statement about reality.

The prime motivation of our studies was rooted in the stark realization of this thesis, namely that much more often than not, objects, on all scales (except for the atomic scale), are *not* symmetric. To appreciate this, one need only refine the resolution of observation – spatial or temporal – to the point where it becomes evident. The advent of highly sensitive analytical and probing tools in modern chemistry shows again and again that even structures which have been classically treated as symmetric, actually are not. Suffice it to consider, for instance, the observation of 'symmetric' molecules on time scales that are faster than typical vibrations or rotation rates; or to consider the local distortive forces on 'symmetric' molecules in the condensed phase, in order to realize this situation. An example of the former is the rotation of the two ethane tetrahedra around the C–C bond. Current wisdom allows an exceedingly poor description of that process from the symmetry point of view: ethane is D_{3d} when staggered, D_{3h} when eclipsed and D_3 anywhere in between. But consider the rotamer which is only 1° away from any of the extremes: is it already a D_3 species? A related phenomenon is the well-known removal of the degeneracy of energy levels of chemical species whenever these are contained in an

environment of symmetry other than their own. The degree of removal of degeneracy is directly linked to the 'decrease' in the symmetry of the environment, compared to that for the isolated chemical species. Traditionally, this problem has been treated in terms of jumps in the symmetry point group.

It appears then that symmetry serves, in many such instances, as an *approximate*, idealized descriptive language of the reality of chemistry and molecular physics. While it is true that an imprecise language helps in grasping complex situations and in identifying first-order trends, the danger of missing the full picture because of a vague description is always awaiting the user of the current symmetry language.

The second motivation for this study is intimately linked to the first one: quite often one can identify in molecules of low symmetry the clear remnants of a higher symmetry. For instance, consider the case of two ethylenes approaching each other for a [2+2] reaction. The answer to the question whether that reaction is allowed thermally or photochemically, or whether a suprafacial or antarafacial process will take place, or whether the reaction will take place at all, is very much dependent on the symmetry of alignment of the two reacting molecules or moieties. The ideal symmetry is D_{2h} for parallel approach, and it is predicted successfully that this symmetry is needed for a suprafacial photochemical formation of cyclobutane. In many cases, however, the two ethylenes are *not* in an ideal D_{2h} alignment because of an intramolecular frozen conformation of the two double bonds, or because of non-symmetric steric hindrance caused by substituents on the double bonds. Yet current wisdom is to treat these cases as if they were in the required ideal symmetry, ignoring the fact that, in these common situations, only a vague memory of the ideal symmetry exists. Or consider the even more elementary case of benzene versus deuteriobenzene: the first has a D_{6h} symmetry, but the latter jumps to C_{2v}; isn't it more natural, from the point of view of the chemical and physical properties, to treat deuteriobenzene as being practically a D_{6h} object to some degree?

These two motivations have led us to propose[4-6] that quite often it is natural to evaluate 'how much' of a given symmetry there is in a structure. Consequently, we have explored the possibility of treating symmetry as a structural property of *continuous* behavior, as complementary to the classical discrete point of view. A continuous symmetry scale should be able to express quantitatively how much of a given symmetry there is in any (distorted) structure, at any temporal resolution, at any spatial resolution, and with reference to any ideal symmetry. Toward this goal, we have designed a general symmetry measurement tool; it is described in the next section. In the sections that follow we discuss some concepts that emerged from our treatment of symmetry and chirality as continuous structural traits.

1.2 The Continuous Symmetry Measure (CSM) Approach

The design of a measurement tool involves some degree of arbitrariness, in the sense that one has to decide on issues such as how should the zero-reference level be set, what should be the maximal value, what should be the actual measurement yardstick, or what normalization procedures should one employ. Such decisions will, of course, be the subject of criticism. Having this in mind, we decided to base our symmetry

measure on a definition which would be as minimalistic as we could practically get. Our answer[4] to the question 'How much of a given symmetry is there in a given structure?' is then: *Find the minimal distances that the vertices of a shape have to undergo in order for the shape to attain the desired symmetry.*

In a formal way, given n vertices of the original configuration, located at P_i, and given a symmetry point group G, the amount, S(G), of this symmetry in this configuration is

$$S'(G) = \frac{1}{n}\sum_{i=1}^{n}\left\|P_i - \hat{P}_i\right\|^2 \qquad (1)$$

where $\hat{P}_i$ are the corresponding points in the nearest G-symmetric configuration. Equation 1 is general and allows one to evaluate the symmetry measure of *any* shape relative to *any* symmetry group or element. S(G) is a 'measure' (a metric). However, in order to avoid size effects, the size of the original structure is normalized to the distance from the center of mass of the structure (placed at the origin) to the farthest vertex, i.e. this distance is 1, and each of the other P_i's is divided by the same maximal P_i. The nearest set of $\hat{P}_i$'s is obtained in terms of the normalized coordinates as well; other normalizations are possible. These features are illustrated in Figure 1. Squared values are taken so that the function is isotropic, continuous, and differentiable. The bounds are $0 \le S'(G) \le 1$. If a shape has the desired symmetry, $S'(G) = 0$. A shape's symmetry measure increases as the shape departs from

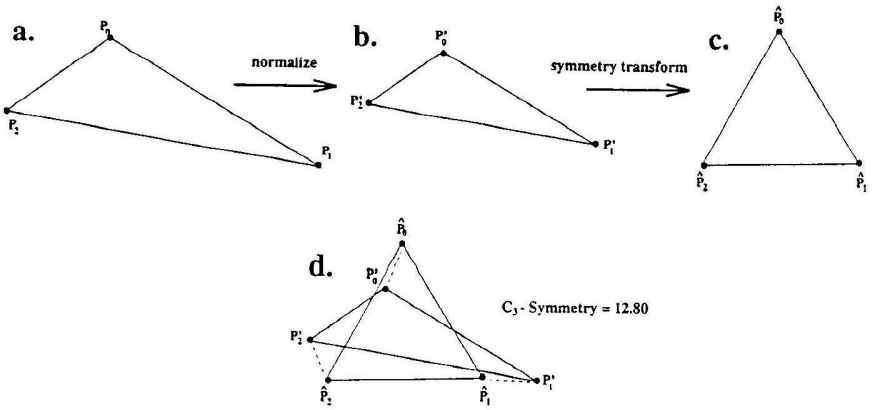


Figure 1 The basic features of the Continuous Symmetry Measure (CSM): In order to evaluate how much C_3-ness there is in the triangle (a), its size is normalized (b), and the C_3 symmetric structure (c) which is nearest to (b) is found using the symmetry transform described in the Appendix. The $S(C_3)$ value is calculated (d) from the minimal distance between (b) and (c), using equation 1. In this case, $S(C_3) = 12.80$.

G-symmetry and it reaches a maximal value (not necessarily 1). The maximal value of 1 is obtained when asking, for instance, how much hexagonality is there in a perfect pentagon? The nearest shape to answer it is a point located at the center, the distance of which from the pentagon vertices is 1. However, in most instances questions of interest are, for example: how much hexagonality is there in a distorted hexagon; how much C_3 is there, and so on; and in such regular cases the maximal possible S′ value is smaller than 1 (see below). Furthermore, it has been our experience that the majority of continuous symmetry related questions focus on distortions which are not extreme, thus yielding small S′ values. We found, therefore, that for convenience it is better to expand the 0 to 1 range by 100:

$$S(G) = 100 \cdot S'(G) \tag{2}$$

All S(G) values, regardless of G, are on the same scale and therefore comparable: one can compare the degree of, say, tetrahedricity of various distorted tetrahedra, the symmetry content of various symmetry subgroups in one tetrahedron (say, its C_3-ness, C_{3v}-ness, etc.), and even different symmetries in different objects. Another important feature of the CSM approach, to which we return again shortly, is that no reference shape is assumed at the beginning of the analysis, though it is obtained as an end outcome.

A particularly interesting family of point groups are the achiral ones, namely all groups which contain improper elements such as reflection, inversion and even-numbered improper rotations. In these cases S(G) is *a measure of chirality,* and one gets then a Continuous Chirality Measure (CCM).[7] In its simplest manifestation, the measure of chirality will be $S(C_s)$, namely the distance of the object from having a symmetry mirror. Because of the central role of chirality in chemistry, many of our examples below are of chirality measurements. We emphasize at this point an important feature which distinguishes our approach from previous propositions to measure chirality: ours yields a quantitative evaluation of chirality, i.e. the distance from improperness, which is a special case of a general treatment of symmetry. Our methodology allows us, therefore, to get a full symmetry profile of a structure: its chirality content and its symmetry content toward any desired symmetry group.

The description of shape through similarity or distance functions is a known approach[8] and so it is in order to emphasize also the following unique aspects of equation 1: the standard approach has been to define a *specific (ideal) reference structure* and to find the distance of the studied structure from this reference. In contrast, we have set ourselves a more general and more demanding task: to find the coordinates of the ideal reference structure (the perfectly symmetric one), namely its shape and size, which are *a priori unknown*. Furthermore, not any reference shape with the desired G symmetry is allowed: one has to search for a specific shape which is the *nearest* from the point of view of equation 1, that is, one that will give the *minimal* S(G). In some cases, the shape of the nearest symmetric object is trivially known (for instance, the nearest C_6 for a distorted hexagon is a perfect hexagon), but more often than not it is unknown (for instance, the specific shape of the nearest C_3

of this distorted hexagon).

The main practical problem was then how to find the nearest symmetric shape to a given one, namely how to locate the specific set of $\hat{P}_i$'s which will minimize S(G), the continuous symmetry measure (CSM). In a recent series of papers,[4-7] a detailed solution to that problem was provided, based on what we termed the 'folding-unfolding' rationale. Despite its key importance to the whole issue, we shall not dwell in this non-technical summary on the mathematical details, the algorithms and the proofs. We illustrate the main features of the folding-unfolding in the Appendix for two elementary cases: rotation symmetry and reflection. The construction of a library of programs that enables computation of various S(G)s is an on-going process in our laboratories; the many examples below and in the cited references illustrate the current possibilities.[9]

Important contributions toward quantitative measurement, in particular to that of chirality, have been made by a distinguished array of researchers, and a bibliography of selected papers is provided for the interested reader in ref. 10.

1.3 Selected Properties of the CSM

Let us now, with the aid of figures and their captions, illustrate some of the properties and potential applications of the CSM.[4-7]

Figure 2 demonstrates that one can take a distorted shape, a hexagon in this case, and answer the following questions. How much C_2-ness does it contain? How much C_3-ness?; and so on. Note that the S(G) values in the caption are comparable to each other, namely, the original shape is closest to achirality and farthest from C_6-ness. Similarly, one can take a series of n-polygons (Figure 3), and see which is closest to its own C_n: it is seen, for instance, that the triangle and the pentagon are similarly distant from their respective C_n's. Finally, one can take this series of polygons, and ask which of them is closest to or farthest from a given symmetry. In this case we determine the $S(\sigma)$ values (Figure 4), namely the two-dimensional chirality, and find that the triangle is the most chiral within this set of specific polygons. In fact, one can also ask what is the most chiral triangle in general? Our methodology points to the triangle shown in Figure 5 as the most chiral one. The most chiral tetrahedron is shown in Figure 6.

The method is not limited to cyclic structures; it can be applied to branched structures (Figure 7), spirals (Figure 8), polyhedrons like the fullerenes (Figure 9), knots and other special topologies (Figure 10), and even to complex structures like random aggregates (Figure 11). The method also allows one to distinguish between global symmetry or chirality of the whole object, and local symmetry or chirality of zones in the structure. An example of this distinction, made for two-dimensional threo and erythro diastereomers, is given in Figure 12.

The changes in symmetry content with mutual three-dimensional orientation of two molecules are shown in Figure 13. This example answers the need, outlined above, of adding gray levels to the strict symmetry demands of 'symmetry dictated' reactions. We are in a position to pose a new type of question for such reactions, which for our example is: what is the relation between the degree of D_{2h}-ness of a

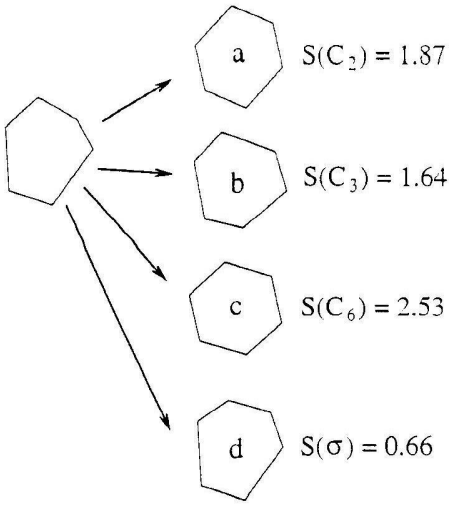


Figure 2 Given a distorted hexagon, one can find its symmetry value with respect to various symmetry elements, as well as how the nearest shape with the desired symmetry looks. (a)-(d) are the nearest C_2, C_3, C_6 and 2D-σ structures. The S values indicate that the hexagon is farthest from perfect hexagonality, and closest to 2D achirality. Note that $S(\sigma)$ measures chirality.

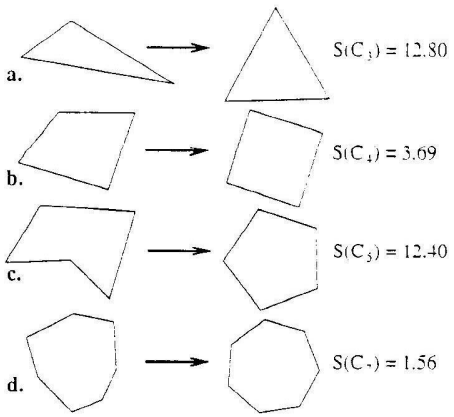


Figure 3 The distance of various n-polygons from their own C_n, and the $S(C_n)$ values.

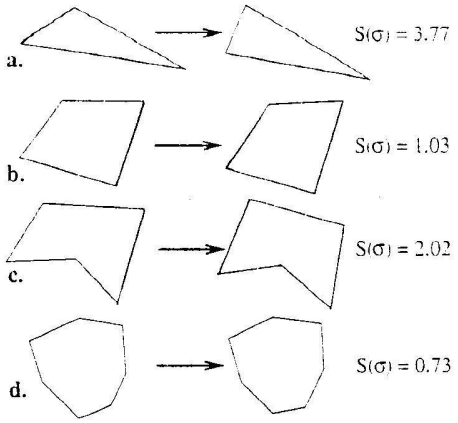


Figure 4 The 2D chirality values of the polygons in Figure 3, and the nearest achiral structures.

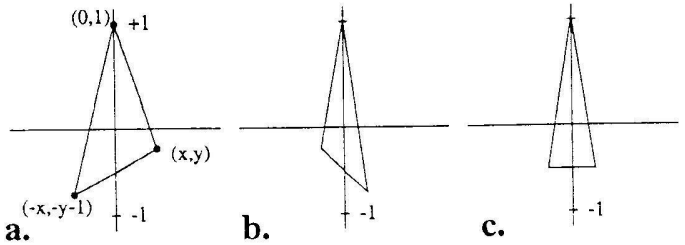


Figure 5 The most chiral triangle (b). It was found by searching over the coordinates shown in (a). The nearest achiral triangle is an isosceles (c). The edge ratio (and the sine of angles ratio) of the most chiral triangle is 1.00 : 0.75 : 0.36, with S(σ) = 3.03.

reacting [2+2] system, and (say) the rate, i.e. allowedness, of the reaction?

Another important feature of the CSM approach is that it can follow the gradual symmetry changes along dynamic processes, such as vibrations, rotations, and conformational changes. Since modern experimental techniques have reached the time scales of these events, symmetry changes within them become relevant. We illustrate this in a number of model systems. The first (Figure 14) follows the $S(C_2)$ content of a water-like molecule along a ν_3 vibration; the second follows the tetrahedricity changes (T_d-ness) during the vibration of a tetrahedron (Figure 15); the third (Figure 16) follows the tetrahedricity content along the pathway of a Walden inversion; and the fourth follows the change in D_{3h}, D_{3d}, and σ (chirality, which, in this case also measures C_{3v}) content in rotating ethane, as a function of the rotation angle (Figure 17). Note that the D_{3h} and D_{3d}, which are traditionally linked with

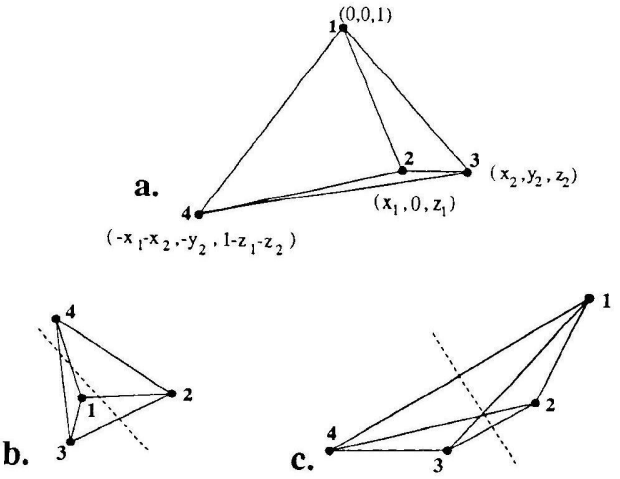


Figure 6 The most chiral tetrahedron was found by searching over all the coordinates shown in (a), and turned out to be a tetrahedron (a simplex) with a C_2 symmetry (b). The C_2 axis is the dashed line bisecting the edges 14 and 23. The nearest achiral structure, in this case (c), is the tetrahedron collapsed to a plane (of this page). The edges ratio of (b) are 1:1:1:1.6:1.6:2.3 for edges 12, 23, 34, 13, 23, 14 respectively, and its $S(\sigma)$ is 4.1.

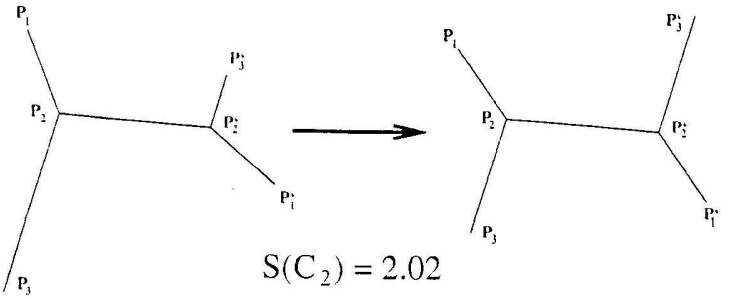


Figure 7 The $S(C_2)$ value and the nearest C_2 symmetric shape of the distorted 'substituted butane'.

ethane, are actually rare events in this molecule, and that in fact, *ethane is chiral* most of the time, with maximal chirality values at $30° + n \cdot 60°$ (Figure 17, top, c). We have termed these rotamers of maximal chirality 'chiramers'.[7]

Probabilistic arguments may be added as well. These reflect the actual status of all experimental techniques for structure determination: the atom locations are given

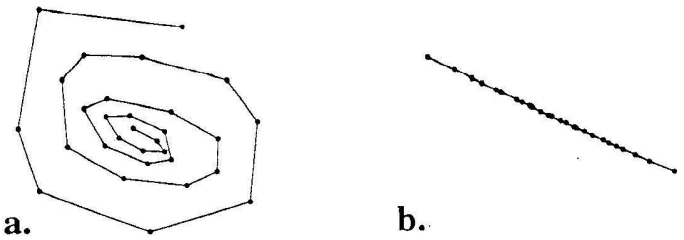


Figure 8 The 2D-chirality value of the spiral (a) is 8.8, and its nearest achiral shape is the mirror line (b).

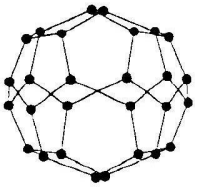


Figure 9 Some of the fullerenes, such as the C_{28} fullerene, are chiral. This fullerene has D_2 chiral symmetry, and its chirality value is 24.9.

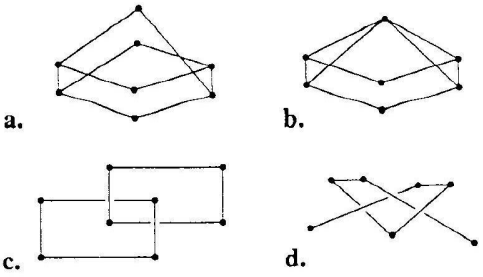


Figure 10 The algorithms of the continuous symmetry measures are capable of handling unusual chiral topologies as well. Shown here are (a) a chiral Möbius structure and its nearest achiral structure (b); a catenane (c) (the two rings are tilted at 25.6° to each other); and a knot (d). The respective $S(\sigma)$ values are 1.95, 0.41 and 0.53.

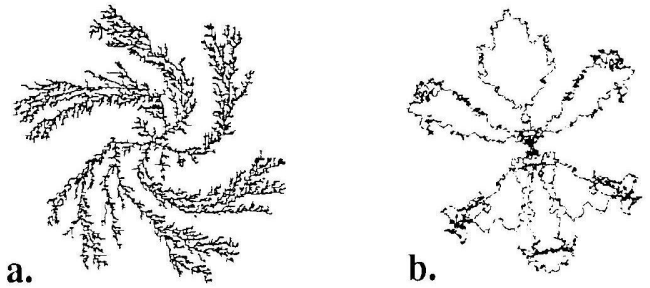


Figure 11 A 2D chiral diffusion limited fractal aggregate (a), and its nearest achiral envelope (b). The chirality value of the envelope of (a) is 3.4.

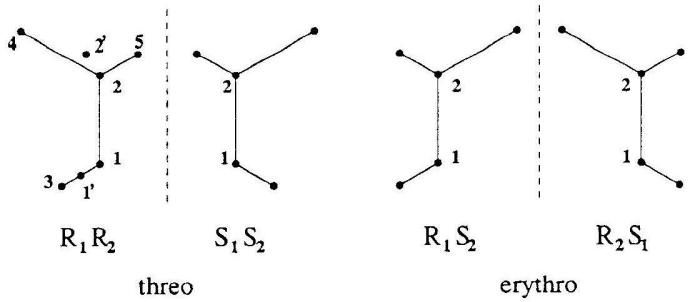


Figure 12 The chirality content of a threo pair must be different than that of the erythro pair (it is 2.5 and 3.2, respectively). While these are the global chirality values, one can also evaluate local chirality of each of the chiral centers, 1 and 2, taking for each the centroids of the other, 1′ and 2′ (shown for threo); the values are 3.4 and 0.29.

only within a certain accuracy (c.f. the thermal ellipsoids in X-ray analyses). For these purposes we studied the following questions:

- What is the closest, most probable symmetric shape represented by data with given uncertainty?
- What is the probability distribution of symmetry measure values for the given uncertain data?
- What is the confidence value in the S(G) value, given an uncertainty in the location of the atoms?
- What is the probability distribution of S(G) under conditions of dynamic randomly changing locations? An example of the latter is given in Figure 18; we return to this issue below when discussing the order in a heated small cluster, and the problematics of the concept of symmetry for large random objects.

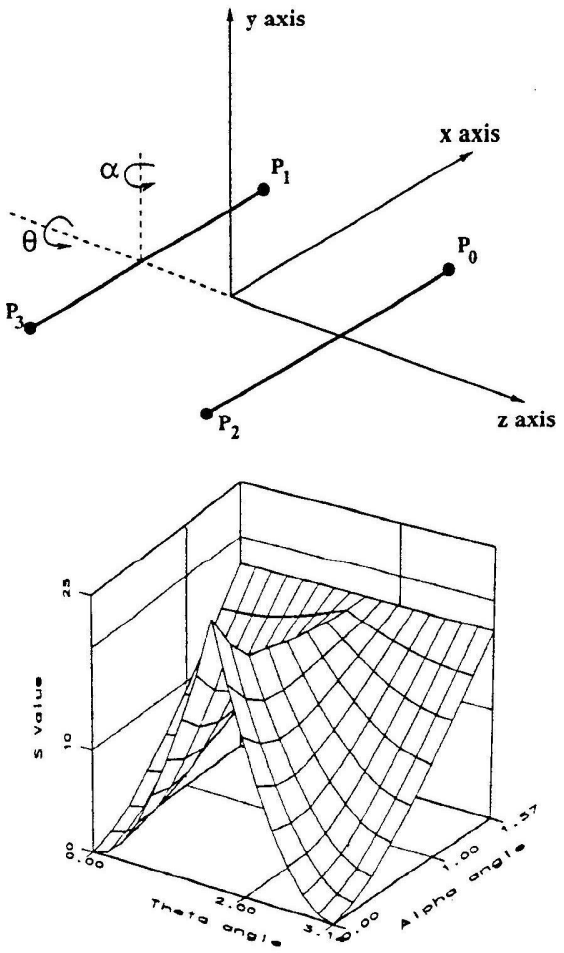


Figure 13 The degree of D_{2h}-ness of two mutually aligned ethylenes (bottom) as a function of two alignment angles (top).

Finally, although this chapter focuses on *geometrical* considerations of symmetry, the approach can be extended to physical parameters, for instance by analyzing the symmetry of equicontours of properties. This is demonstrated in Figure 19, details of which are given in ref. 7.

2. Concepts II - IV: Symmetry as a Process Coordinate; Non-Handed Chirality; Isochirality

An inherent property of enantiomerization pathways is that these need not pass through an achiral intermediate.[11] Considering that imposing a reflection plane on a

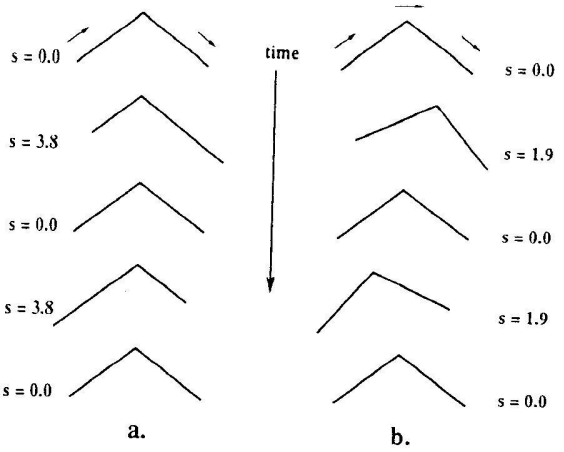


Figure 14 The degree of C_2-ness in a vibrating ABA molecule at various snapshots during one vibration cycle: (a) the vibration is of ν_3 mode; (b) as in (a), but B is anchored to a 'surface'.

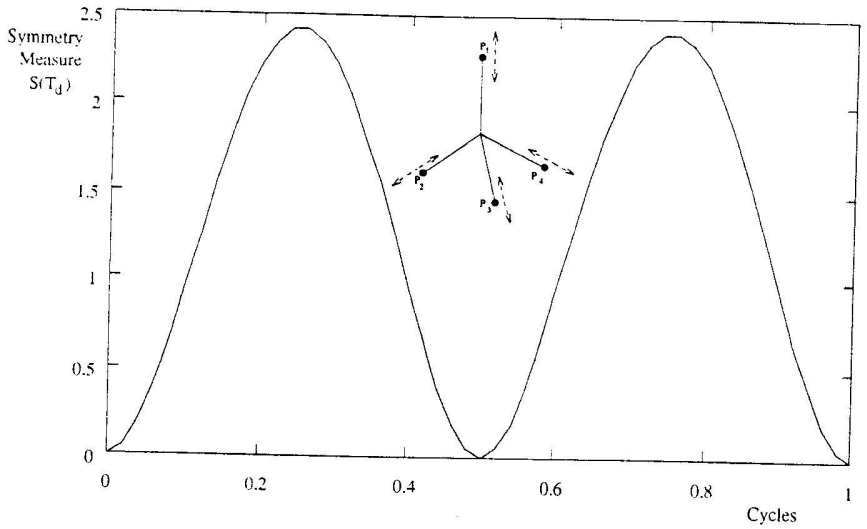


Figure 15 The variations in the T_d symmetry content of a vibrating tetrahedron during one cycle. (In this model the vibration extends the arms to 1.25 of the remaining length, and the two pairs of arms vibrate with a phase delay of 180°.)

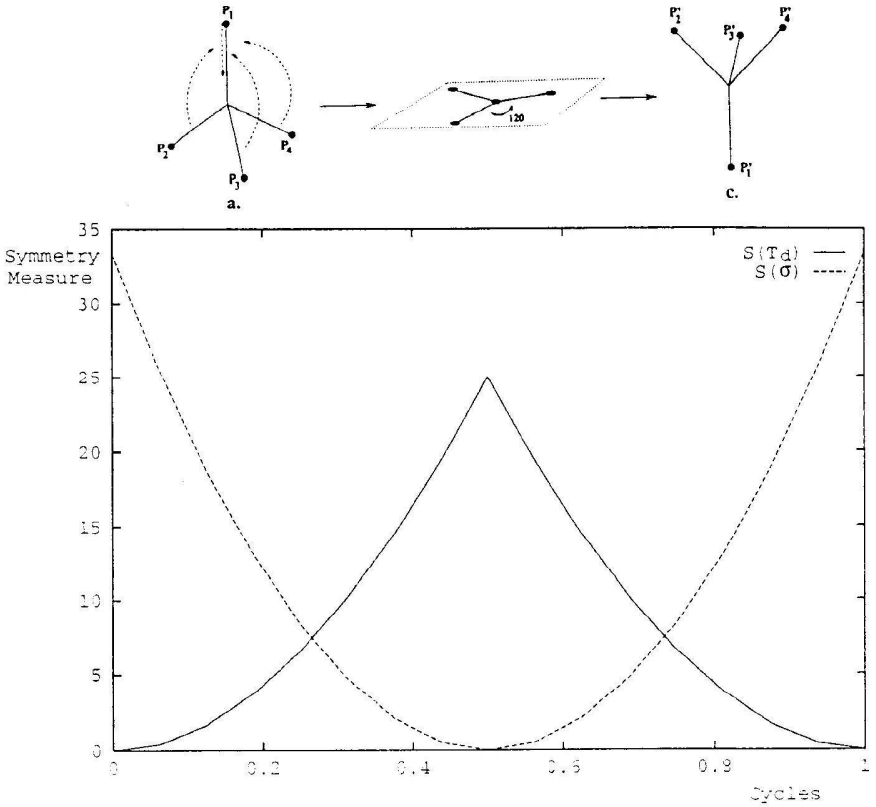


Figure 16 The variations in $S(T_d)$ (full line) and $S(\sigma)$ (dashed line) of a tetrahedron (bottom) along a Walden inversion (top).

reacting molecule that *a priori* does not possess it is a rather restrictive demand; one realizes that achiral enantiomerization pathways (i.e. ones that pass through an achiral intermediate structure) should be an exception.

The CSM approach allows us to study this property on a quantitative level. Figure 20 is an illustration of the phenomenon: two enantiomerization pathways are shown for the same structure, one of which passes through $S(\sigma) = 0$, and the other never reaches zero level. Chiral pathways carry an interesting feature: Let us define the structure on the left in Figure 20 as Left and the one on the right as Right. Somewhere along the path from Left to Right there is a transition point between the properties of Leftness and Rightness, namely a chiral structure for which handedness cannot be assigned under the specific definitions used for this case. We have termed this situation *non-handed chirality*.[12]

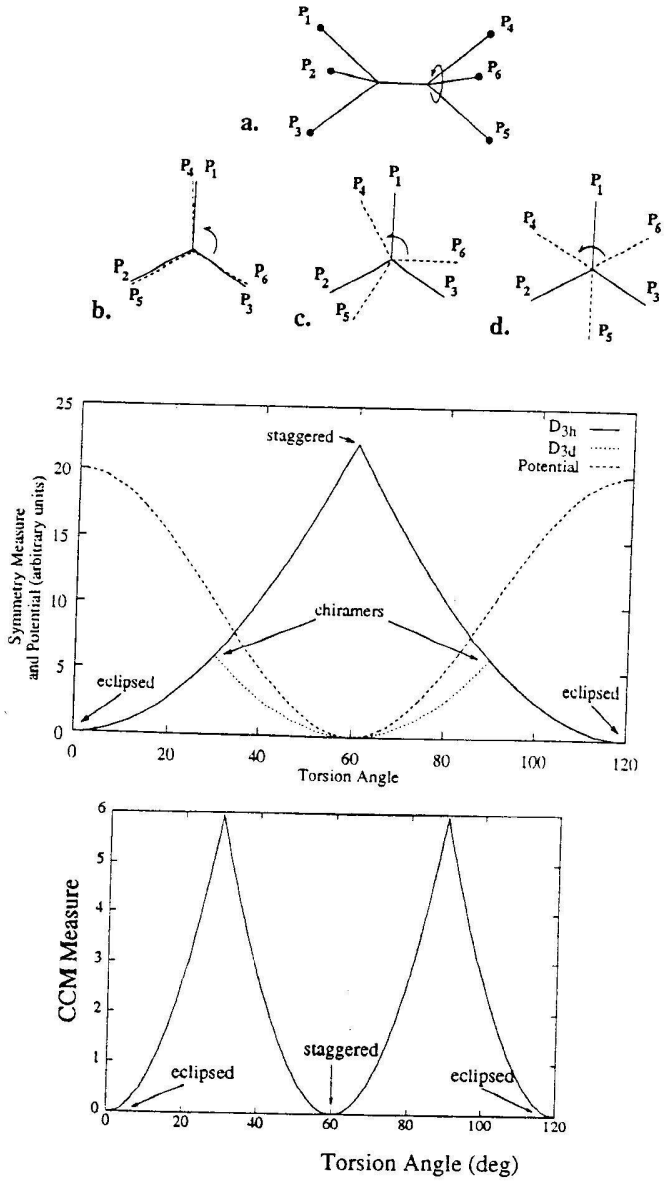


Figure 17 Symmetry variations in rotating ethane (top, a). Shown are the variations in D_{3d}, D_{3h} and C_{3v} (middle), and chirality (σ or C_{3v}, bottom). Note the locations of the eclipsed (top, b), staggered (top, d), and the most chiral rotamer, the chiramer (top, c). Also shown are the sinusoidal variatiaons in the potential.

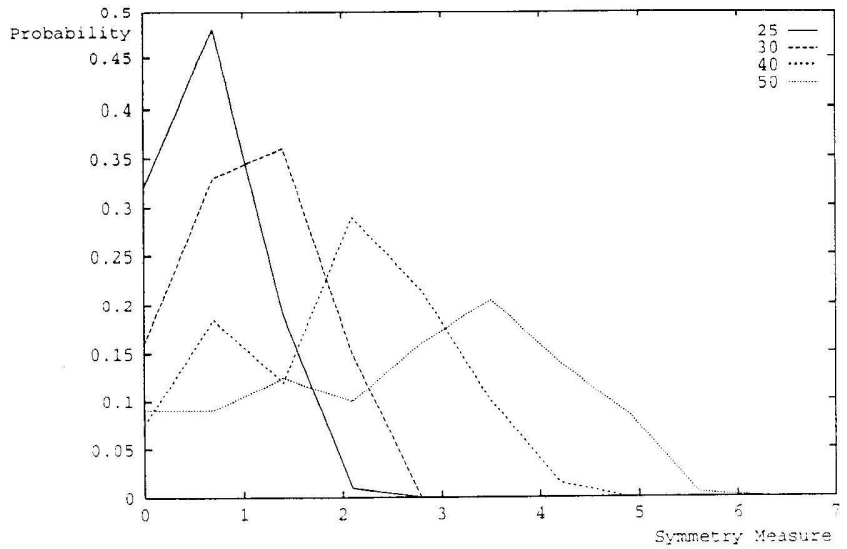


Figure 18 A tetrahedron vibrates with random phase shift between the four arms (as in a central heavy atom linked to 4 light ligands). Shown is the probability distribution of having a given tetrahedricity, $S(T_d)$, for four cases of maximum arm extension, 25, 30, 40 and 50% of the original length.

Definitions

Handedness. The arbitrarily assigned structural property of leftness/rightness. (The relevant definition of handedness for structures along a reaction pathway is usually that of the reactants and/or products.)

Non-handed chirality. Chirality for which it is not possible to assign handedness under a given definition of this property. *Non-handedness* is the property of this impossibility. (*Achirality* is trivially non-handed.)

Non-handed structure. A structure, the chirality of which is non-handed.

Point of non-handedness. A point along a reaction pathway representing a non-handed structure.

Non-handedness is intimately linked to the question of the chirality of Ruch's potato (Ruch in ref. 10). In this context we make the following claim. Select a definition of handedness; since for a potato (or, say, a cluster) there is an infinite number of chiral enantiomerization pathways, there is an infinite number of non-handed potatoes, under the selected definition, located on these pathways. Since this is true for any selected definition, one is bound to conclude that it is impossible to find a general definition for handedness that encompasses all potatoes. The same reasoning holds, in fact, for any racemizable molecule (provided that it has at least five atoms).[11d,e]

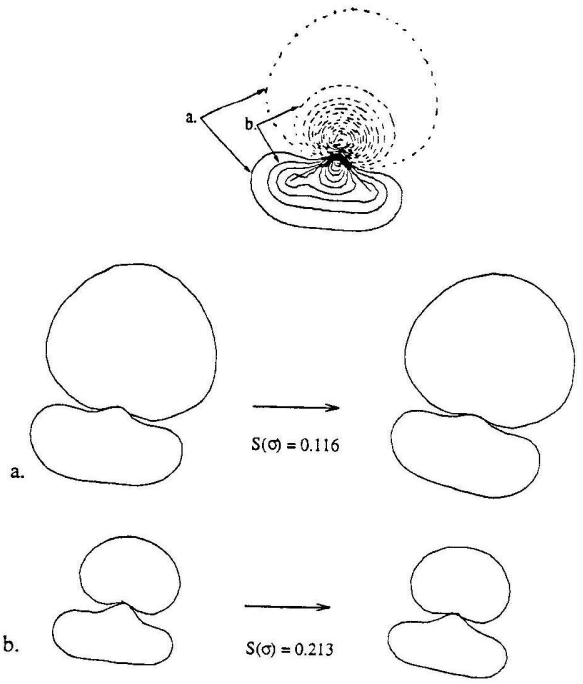


Figure 19 Two equi-amplitude contours of the wave function of the lone pair orbital of a distorted water molecule. Shown are the 2D-S(σ) values for the outer contours. The value for the next inner one is 0.248.

We illustrate now some practical implications of this discussion on the chirality of the cyclic trimer of water.[12] The existence of a water trimer, which was predicted by Scheraga et al.[13] in the early seventies, has been observed recently by Pugliano and Saykally[14] using far infrared (vibration/rotation) tunneling spectroscopy. The observation has triggered intense theoretical and experimental activity on the structure and energetics of this and other small clusters of water molecules.[15] A remarkable structural property of the trimer obtained by virtually all the computational methods employed is that the minimum energy structure of the trimer is chiral (Figure 21). By labeling each of the six hydrogens and the three lone pairs of the trimer, and by employing the C_{2v} symmetry of H_2O, 96 isoenergetic trimer minima exist in the multidimensional potential surface, comprising 48 enantiomeric pairs, such that neighboring minima are chiral pairs.[13] Motion from one minimum to another thus leads either to conversion of the handedness (enantiomerization) or to its retention (automerization). Our methodology allows one to explore quantitatively the continuous changes of the chirality value along selected interconversion pathways, namely to *use the symmetry measure as a process coordinate.*

As seen in Figure 21, the trimer has two hydrogens on one side of the oxygen

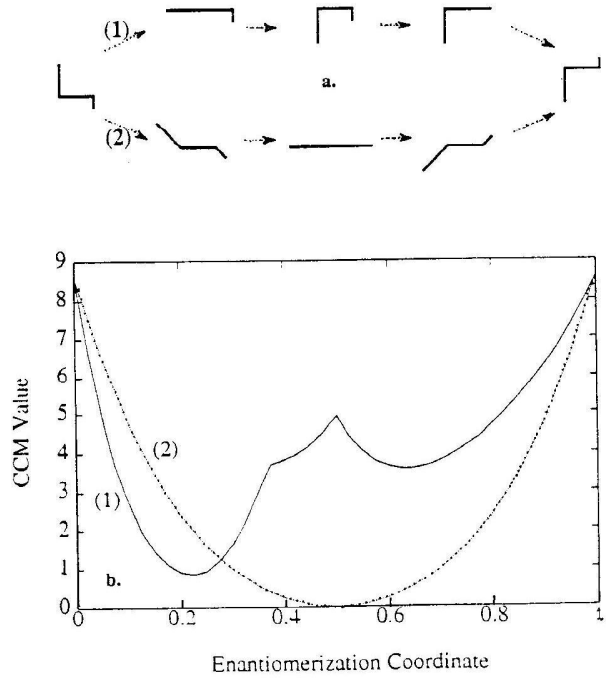


Figure 20 Enantiomerization pathways (a) can be either achiral (2) or chiral (1). In the latter, the continuous chirality measure (CCM) never drops to zero (b). (The process coordinate is the sum of angle changes.)

plane ('above') and one on the other ('below'), and alternating in-plane O–H and H···O bonds. Enantiomeric interconversion is therefore possible by mechanisms which either move one hydrogen from above to below or which change the direction of alternation of the oxygen-hydrogen bonds. These two mechanisms known as the *flip* mechanism and the *clockwise/counterclockwise (cw-ccw) rotation* mechanism are shown in Figure 22. (For other mechanisms see our full report in ref. 12.) As seen in Figure 23, the flip pathway is all-chiral, and, interestingly, with a chirality value which remains nearly constant throughout the enantiomerization. As calculations showed that this hydrogen flipping mechanism is energetically the most facile,[13,15] it is intriguing to consider the role of the constant $S(\sigma)$ value in the context of the very low free-energy barrier for that process: a constant chirality value may minimize the entropy change for a process.

The transition state for the flip mechanism is obtained when the flipped hydrogen is approximately within the average oxygens' plane (Figure 22a). As the enantiomerization pathway is chiral, the transition state must be chiral as well.[15] Let us now link it to non-handedness, taking the chirality definition of handedness in[14]. Somewhat reformulated, the definition is as follows (Figure 21): with the projection

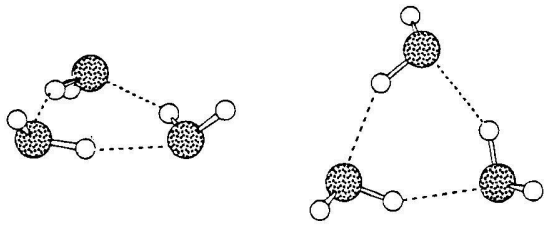


Figure 21 The handedness of the chiral water trimer, (a), is defined by watching the projection of the trimer from the side of the oxygens' plane which contains the two 'free' hydrogen atoms (b). This side of the cluster is called 'above' the oxygens' plane. If the three in-plane O–H bonds are arranged counterclockwise, as in (b), it is left-handed, and if they are arranged clockwise, it is right-handed.

of the trimer from the side of the oxygens' plane which contains more 'free' hydrogens (above), if the three in-plane O–Hs are arranged counterclockwise, this is left (it is a left spiral if one follows accordingly the three non-bonded hydrogens). Following this chirality convention, we find the non-handed structure in close proximity to the transition state:[16] at the transition state, represented by the middle structure in Figure 22a, one of the two formerly 'above' hydrogens is nearly in-plane,[16] where the definition of the trimer handedness collapses. Since for this structure there are no 'above' and 'below' plane-sides, which are used to define right and left, it is non-handed under our definition, yet it is chiral: the pair of its enantiomers is shown in Figure 24. It is now possible to take this new enantiomeric non-handed pair and provide a new left/right assignment, using a different definition of handedness. However, such a definition would fail again in an all-chiral enantiomerization of these structures; and so on.

There are, in fact, three possibilities for the flip mechanism (each of the three hydrogens), and Figure 25 shows them all. It is seen that the pathways for the flipping of the two 'above' hydrogens are mirror images of each other. The pathway itself is not mirror-symmetric around the transition state, because of its C_1 symmetry (the two 'above' hydrogens are unequal in their immediate neighborhood, Figure 21). Whereas the enantiomeric pair of pathways (a,b in Figure 25) is obtained by flipping two different 'above' hydrogens, there is another possibility of obtaining a mirror-image pathway, this time from reverse flipping of the same hydrogen: following the arguments of Wolfe and Salem.[11b,c] For each chiral pathway, there must exist also its exact mirror-image pathway, because for each chiral conformation along the path, an equi-energetic enantiomer exists. A third flip pathway (c in Figure 25) is the flip of the lone hydrogen from 'below'. This one connects two distinctly different isomers which are not enantiomers of each other, as is clearly seen at the

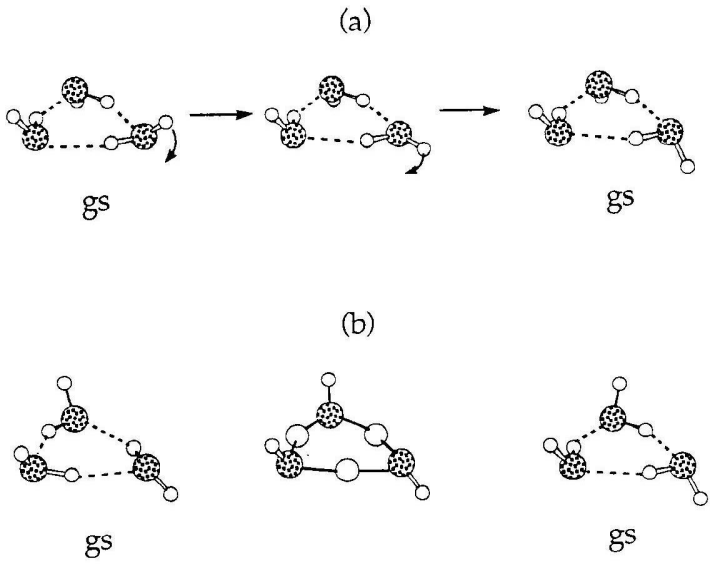


Figure 22 Enantiomerization pathways of the trimer of water: (a) Flipping is accomplished by rotating one water molecule about its in-plane O–H bond. This leads to enantiomerization if the free hydrogen of the monomer is one of the 'above' plane hydrogens. (b) Clockwise or counterclockwise (cw-ccw) motion is a synchronous rotation of all six hydrogens. The three in-plane hydrogens exchange hydrogen bonding with covalent ones. Note the C_s symmetry of the transition state (middle structure). gs: ground state.

end points of the curve: the right-hand side of the curve is of the 'C_3 bowl' trimer shown in the insert in Figure 25. This pathway demonstrates the ability of our approach to track quantitatively chirality changes, not only in enantiomerizations but in practically any isomerization.

It is interesting to consider the intersection points of the various pathways: they are, in general, *isochiral isomers*. *Isochirality* is a concept which is possible only in conjunction with the notion that chirality is a measurable structural property – the central theme of our studies. Whereas for the non-intersecting points (1), (2) and (3) in Figure 25 it is clear that the two isochiral structures are neither identical nor enantiomeric, this is true also for the intersection points shown in the Figure, except for the center intersection point of two enantiomeric pathways, (5), which represents the isochirality of two enantiomers. Isochirality does not imply in any way iso-handedness, although intersection points such as (5) can be points of non-handedness.

To complete the picture, we now comment on the achiral clockwise or counterclockwise (cw-ccw) rotation mechanism (Figures 22b, 23). Achiral pathways should be, by and large, rare in inversion processes. For instance, in the present

example the passage of the cw-ccw route through achirality is intimately linked to the synchronicity in motion of the six nuclei involved. Most random deviations from synchronicity will lift up the minimum from achirality (S = 0). Synchronicity is a rather strict condition, demanding, for instance, high symmetry, which is not the case for the C_1 trimer: the O···H bonds, which are unequal from the point of view of their neighboring atoms, are therefore not expected to move synchronously. As for non-handedness in the cw-ccw pathway, here the transition state is achiral and therefore trivially non-handed. We are currently exploring the proposition that transition states of enantiomerization pathways are in general in the vicinity of non-handed structures.

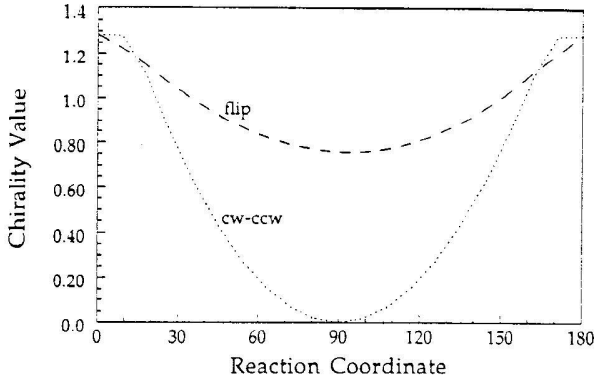


Figure 23 Chirality changes along two reaction coordinates (divided arbitrarily from 0 to 180). Note that the achiral cw-ccw path reaches a minimum of zero chirality at the middle of the pathway. The chiral flip pathway has an almost constant chirality value along the path, with a shallow minimum.

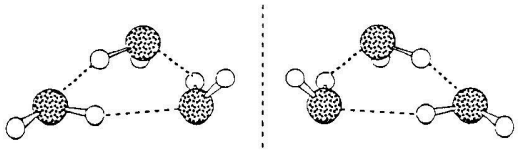


Figure 24 The enantiomeric pair of non-handed structures of the chiral water trimer (see text). They are in the vicinity of the transition state of the flip reaction.

3. Concepts V - VII, Large Random Objects: Incidental vs Inherent Chirality; Virtual vs Natural Enantiomers; Symmetry as a Resolution-Dependent Property

The current intensive interest in disordered and semi-ordered structures (polymers and oligomers, aggregates, clusters, liquid crystals and other supramolecular

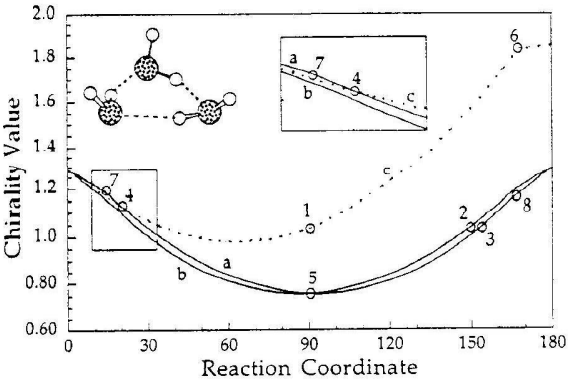


Figure 25 Three flip routes of the water trimer. (a) and (b) are flips of the two 'above' hydrogens. These paths are mirror images of each other chirality-wise and energy-wise. (c) is the flip of the 'below' hydrogen. This flip results in the isomer of the water trimer, shown in the insert. (1), (2) and (3) are isochiral isomers. (4) is an intersection point that represents isochiral isomers of two different paths. (5) is the middle point of two enantiomeric paths ((a) and (b)). It represents two non-handed enantiomers (trivially isochiral). Point (6) is non-differentiable due to a change of the nearest σ plane. Points (7) and (8) are non-differentiable because of a normalization vector flip from the center to the farthest atom.

structures), many of which are chiral,[17] led us to re-explore some conceptual aspects of chirality associated with such structures. We begin by showing that while the classical definition of chirality is well suited to deal with small molecules, it fails in the accurate description of large random structures.

We concentrate here on aggregate structures of the type shown in Figure 26a, known in short as DLAs (Diffusion Limited Aggregates). These arise from a very wide variety of both growth phenomena (polymerizations, coagulations, electrodepositions, surface island formations, etc.) and disintegration phenomena (dissolution, forced penetration, etc.).[18] Although based on entirely random (Brownian) diffusion pathways of building units, the DLA in Figure 26a is chiral under the classical definition: it does not coincide with its two-dimensional (2D) mirror image (mirror line, in 2D), shown in Figure 26b. The non-triviality of associating chirality with a random object (cf. Ruch's potato,[10] Section 2, and Mislow's discussion of the chirality of large ensembles[19]) leads us to create DLAs in which chirality is *inherent*, by biasing the diffusional pathway of the aggregating particles from purely Brownian to chiral. For instance, if clockwise diffusion is (partially) favored over diffusion in the counterclockwise direction, a chiral object reflecting this bias will form. A typical result[20] is shown in Figure 27a.

The DLAs shown in Figures 26 and 27 raise the following problems, not encountered in small molecules:

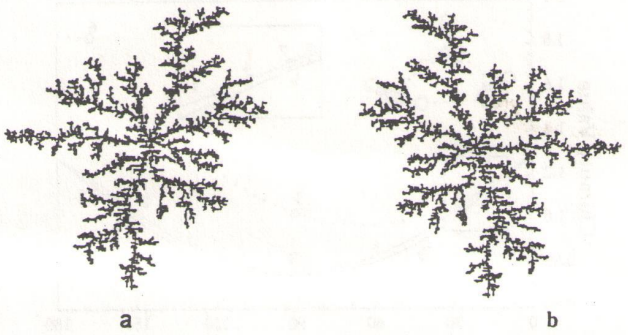


Figure 26 (a) A fractal diffusion limited aggregate (DLA) and (b) its mirror image. The DLA (a) is *incidentally* chiral, and its enantiomer (b) is a *virtual* one.

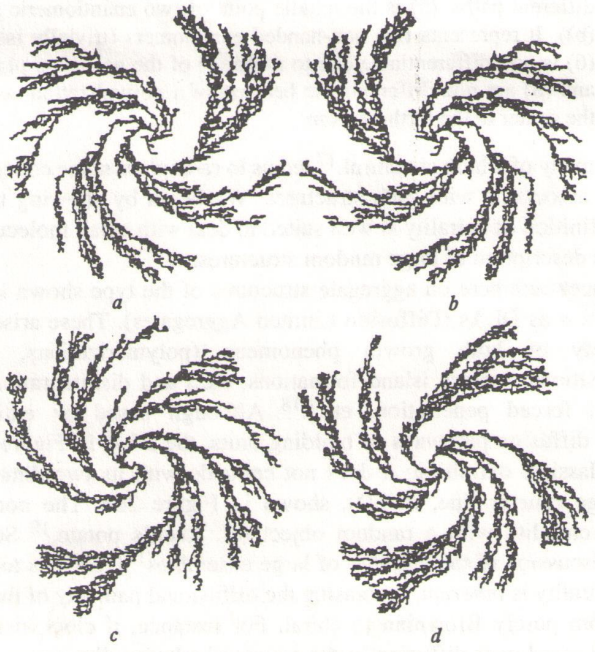


Figure 27 (a) One way of obtaining chiral DLAs is to take the DLA in Figure 26 and rotate it clockwise around its centre, against friction; we assign therefore the symbol *R* to it. (b) The virtual *S* enantiomer of (a). (c) and (d) are two *natural S* enantiomers of (a).

Problem 1. What is the enantiomer of the chiral DLA shown in Figure 27a? Figure 27b shows an artificial reflection of 27a. The problem is that, being a random process, not only can 27a never be repeated, but the enantiomer 27b can never form. What then is the meaning of chirality of an object that in practice can never have an exact enantiomer?

Problem 2. For the sake of further discussion, let us call the reflected shape of the random object the *virtual enantiomer* (Figures 26b and 27b); and the *natural enantiomer* the object obtained by the same random process, only biased in the opposite (enantiomeric) direction, as seen in Figure 27c. There is an *infinite number* of natural left-handed enantiomers, (Figures 27c,d) for each right-handed structure (Figure 27a). What then is the meaning of an enantiomeric pair, when each part of this pair has an infinite number of distinctly different structures; and, taking a single member of the infinite group of clockwise (*R*) enantiomers, how justified is it to pick any one of the infinite counterclockwise (*S*) structures, and call both an enantiomeric pair, a pair which is definitely not a reflection of the other?

Problem 3. Coming back to the DLA in Figure 26a and to the fact that it is chiral because it is random, how can one differentiate between *incidental chirality* (Figure 26a) and *inherent chirality* (Figures 27a,c,d)? What is it that makes incidental chirality distinctly different at least by intuition from inherent chirality? Is there any reason to make that distinction? Should one refer to objects that are incidentally chiral (almost any macroscopic object around us) in terms of chirality at all?

Problem 4. We termed structures 27c and 27d *natural enantiomers* of 27a. Since their structures are different in detail, how can one assess *quantitatively* whether the *chirality content* of structures 27b-27d is similar? Similarly, Figure 28 shows a series of DLAs with a clearly changing spirality character. How can this be measured?

Problem 5. The building blocks of the chiral DLAs are not chiral (square pixels, in our case); yet the whole structure is. (For experimental observations, see ref. 21.) Where is the transition from achirality to chirality? What is the minimal size of features that contribute to the chirality of the whole? (The equivalent question for small molecules is to ask, for instance, where the chiral center is.)

Problem 6. Whenever large objects with many details are analyzed, the question of resolution of the measurement is obvious. Thus, whereas the chirality of the equilibrium structure of, say, 2-iodobutane is uniquely defined because all information on all the positions of the nuclei is given quite accurately, the situation with, say, proteins is different: their structures can be determined at various degrees of resolution. Our sixth problem, therefore, is what is the *resolution-dependency of chirality*? This problem (cf. Mezey's treatment of resolution based chirality and similarity measures)[10,22] is intimately linked to the problem of the fractal nature of the chiral DLAs, described next.

Problem 7. A central structural property of many artificial and natural random objects, which has been at the center of intensive studies in the past fifteen years, is that many of their properties scale isotropically with size as a power law, i.e. they are apparently fractal objects.[15] For instance, in DLAs the mass, m, scales with the

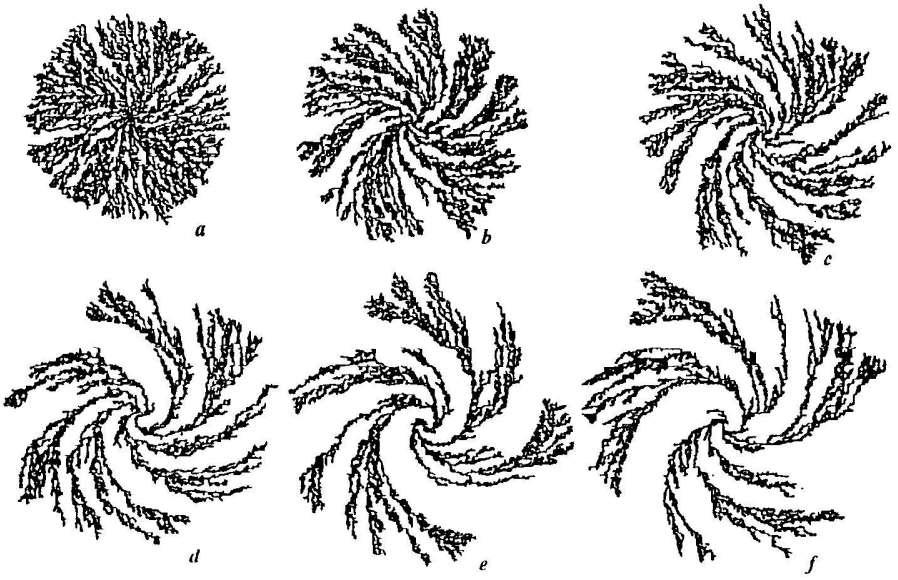


Figure 28 DLAs of varying degree of chirality. (The variations are achieved through the construction parameters of the structure (see text)). Here $N = 10^4$, $p = 1.0$ and $\Delta r = 0.2$ were kept constant, while Δa was varied as follows: a: 10, b: 22, c: 30, d: 50, e: 70, f: 90.

radius, r, where D is the mass fractal dimension:

$$m \sim r^{-D} \tag{3}$$

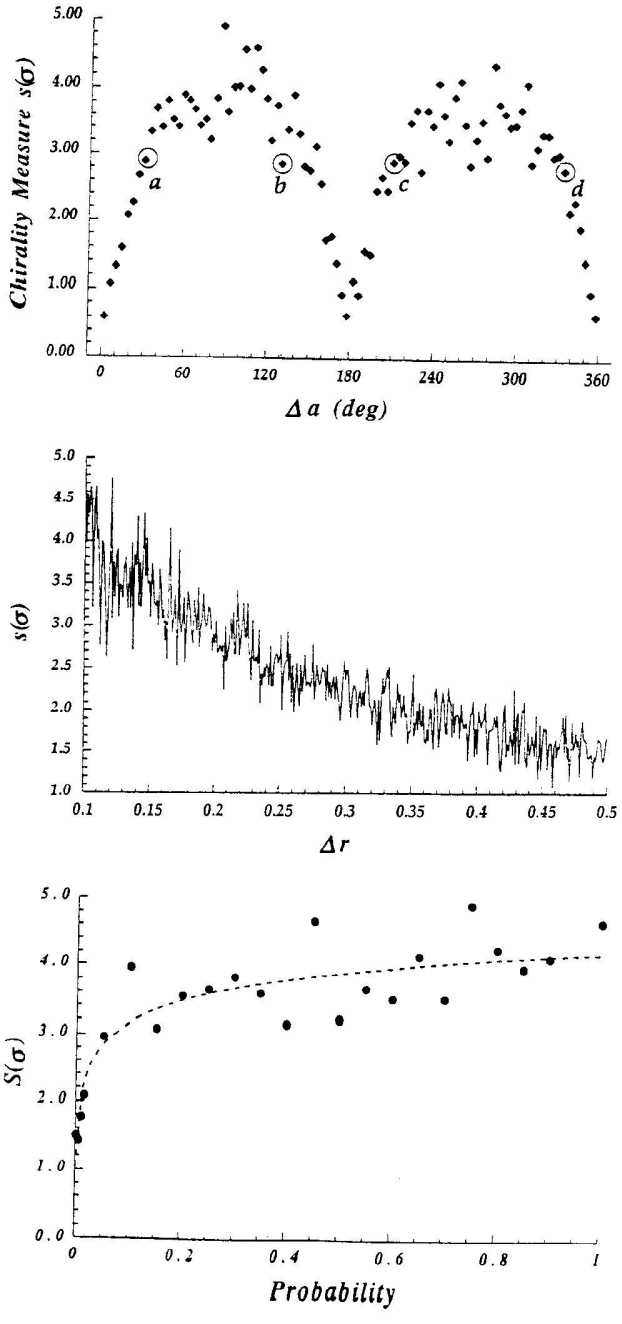
It has been shown that an approximate geometric interpretation of equation 3 is scale or resolution invariance, that is, a large DLA (with r_1) and a small DLA (with r_2) are indistinguishable if the large DLA is reduced in size from r_1 to r_2. Having this in mind, let us now cite Prelog's definition of chirality:[23] 'An object is chiral if it cannot be brought into concurrence with its mirror image by translation and rotation.' The last two are symmetry operations; but so is the change of scale for self-similar objects, where the symmetry operation is dilation/contraction. Shouldn't Prelog's definition be extended to 'translation, rotation and scaling'? Are objects 27a and 11a, which differ only in size (10^4 and 22,880 particles, respectively) but otherwise are constructed with the same parameters, an enantiomeric pair?

To explore these questions, the continuous chirality approach was implemented on the DLAs.[20] DLAs were built, as mentioned above, with one particle after the other randomly diffusing to a growing cluster, and sticking to it upon collision. This

procedure results in *incidental* chirality (Figure 26), and *inherent* chirality is induced in such structures if the diffusion of the random walker is chirally biased, which, in our application, is a preferential probability to move in a clockwise or counterclockwise motion along a spiral direction in each step of the random walk. The spiral we selected is a logarithmic one. For computational convenience we built this spiral by a recursive procedure, in which motion along a spiral line from x_n, y_n, to x_{n+1}, y_{n+1} is achieved by changing the radius with an incremental radial translation, Δr, and by an angular translation, Δa. The shape of the spiral is determined by the values of Δa, Δr. Basically, as Δr or Δa increases (up to ~90° for the latter), the spiral spreads out faster. Similarly, the shape of the chiral DLA is dictated both by the shape parameters of the underlying spiral direction (Δr, Δa) and also by the probability, p, to move in that direction. Figure 28 shows how chirality becomes more pronounced in a series of DLAs built from $N = 10^4$ particles, by changing Δa while keeping all other parameters fixed ($p = 1.0$; $\Delta r = 0.2$). Our methodology allows one to express this visual increase in chirality on a quantitative level, as shown in Figure 29a where $S(\sigma)$ is plotted as a function of Δa. See the figure caption for details. Figures 29b,c show how the chirality content of the DLAs is controlled by two other structural parameters. Figure 29b shows the relation between $S(\sigma)$ and Δr: as Δr increases, the curling of the spiral opens, and the chirality of the DLA decreases. Figure 29c shows the effect of changing the magnitude of the probability, p, to move along the spiral path: as expected, $S(\sigma)$ increases with p, showing high sensitivity to p changes at low p values.

The random nature of the construction of the DLAs and the resulting noisy behavior of Figures 28 and 29 provide answers to some of the questions raised above. It is clear that although the computation of an $S(\sigma)$ value of a specific DLA is unique, understanding and evaluation of the $S(\sigma)$ are possible within the determination of a statistical distribution of $S(\sigma)$ values for a given set of parameters. The question is then: are the $S(\sigma)$ values of inherent chiral DLAs significantly higher than those obtained for incidental chiral objects? And can this type of analysis be a distinguishing tool between these two essentially different types of chirality? Figure 30 shows the frequency of $S(\sigma)$ values for incidentally chiral DLAs and for two sets of inherently chiral ones; bell-shaped, relatively narrow, normal distributions (the parameters of which are given in the caption of Figure 30) are obtained. In these, the inherently chiral DLAs are significantly outside the region of incidental chirality. Obviously, given other construction parameters, the normal distribution bells may overlap in part, and the distinction between incidental and inherent chirality may become statistically blurred.

Next we analyze properties of the DLAs associated with resolution and scale. In this context, how does the chirality content of a DLA change with the history of its growth, namely with the number of particles that build it? The result (Figure 31), surprising at first glance, is that $S(\sigma)$ remains fairly constant from the initiation of the DLA (N is around 1.5×10^3 particles) up to $N = 10^4$. This is actually a manifestation of the self-similar fractality of these aggregates: brought to the same scale, small and large DLAs are indistinguishable from the point of view of their chirality. This scale


Chirality Measure s(σ)
5.00
4.00
3.00
2.00
1.00
0.00
a
b
c
d
0
60
120
180
240
300
360
Δa (deg)
s(σ)
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Δr
S(σ)
5.0
4.0
3.0
2.0
1.0
0.0
0
0.2
0.4
0.6
0.8
1
Probability

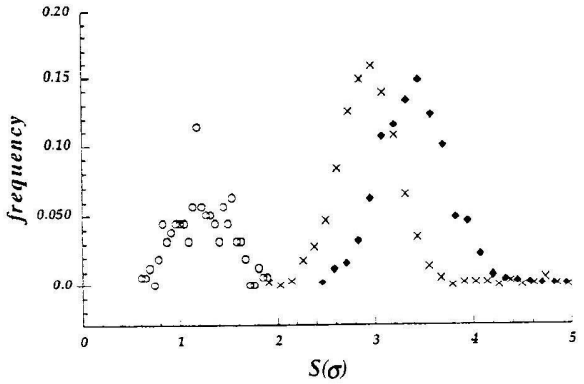


Figure 30 The distribution of S(σ) values for a population of incidental DLAs, (o), and for two sets (x, ♦) of inherently chiral DLAs. o: Number of DLAs (n_D) = 158; x: n_D = 900, Δa = 30°, Δr = 0.2, p = 1.0; ♦: n_D = 947, Δa = 139°, Δr = 0.2, p = 1.0. The bell-shape parameters, from left to right, are: S(σ) = 1.23 ±0.27; 2.96 ±0.37; 3.43 ±0.43.

invariance is also evident in Figure 32, where a set of DLAs of varying size is analyzed according to equation 3. The scale invariance, both to global structure and to the chirality, is a unique feature of chiral DLAs.

We are now in a position to propose some answers to the problems raised above:

1. For large, random chiral objects, an enantiomeric pair has significance and meaning within a statistically large collection of objects.
2. Each and every member of this collection is chiral in itself, and there are two

Figure 29 ◄— The dependence of the degree of chirality of DLAs, on various structural parameters. *Top*: Quantitative evaluation of the degree of chirality of the fractal DLAs in Figure 28, and the relation between the 2D chirality measure, S(σ), and the structural parameter Δa. It is seen that the chirality value increases with Δa up to a maximum value at Δa = 90° where the spiral curvature is maximal, beyond which it drops to a minimum at 180°, and then the cycle repeats itself at the 180-360° interval. The minima do not reach zero, but a small chirality value typical of incidental chirality. Except for the extrema, there are four *isochiral* DLAs (DLAs with the same S(σ) value) in one full cycle of 360°: **a** and **b** are separated by (π – 2Δa), and are structures of the same handedness (homochiral); so is the pair **c**, **d**. The pairs {**a**,**b**} and {**c**,**d**}, separated by π, are *natural* enantiomers of each other: **a** is an enantiomer (and not *the* enantiomer, as is the language for small molecules) of **c** and/or **d**. *Middle*: The dependence of S(σ) on the structural parameter Δr ($N = 10^4$, Δa = 30°, p = 0.7). *Bottom*: The dependence of S(σ) on the probability, p, to walk along the spiral line ($N = 10^4$, Δa = 90°, Δr = 0.2).

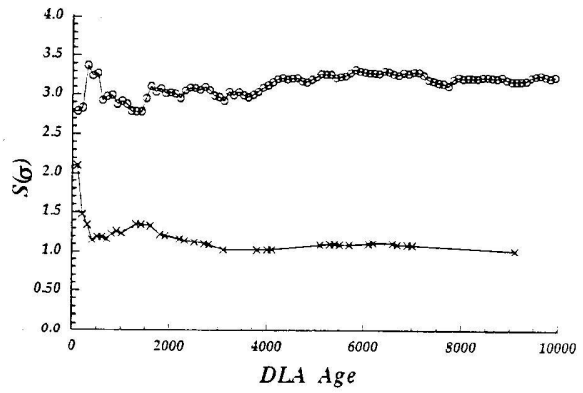


Figure 31 The S(σ) values of an inherently chiral DLA (top) and an incidental chiral DLA (bottom), as the DLAs grow. (Top: $\Delta a = 139°$, $\Delta r = 0.2$, $p = 1.0$; bottom: same but with $p = 0$.) 'Age' refers to the number of particles.

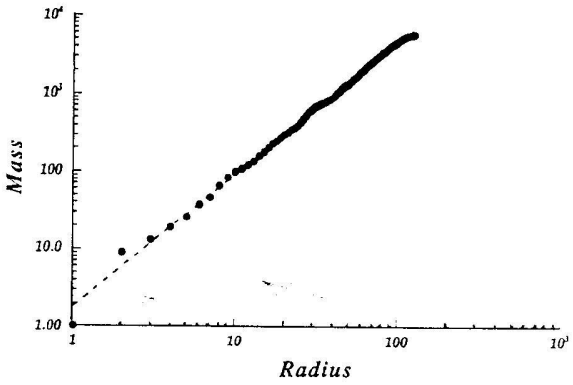


Figure 32 Mass scaling fractal analysis of the set of iniherent DLAs in Figure 31. The apparent fractal dimension is 1.7. (Parameters: $N = 10^4$, $\Delta a = 90°$, $\Delta r = 0.2$, $p = 1.0$. Mass: number of pixels; radius in pixels.)

types of answers as to what is its enantiomer. First, there is the enantiomer which is obtained by perfect reflection. As this enantiomer can be obtained *only* as a mirror image and never in a real repetition of the actual (chemical) process by which the original chiral object was obtained, we propose to term the formal structure a *virtual* enantiomer. Then there is a *natural* enantiomer, which is any member of the collection of objects obtained by repeating the construction in an enantiomeric way.

3. The last distinction must also be made with the collection of structures of the same handedness (for instance, the direction of spirality in our case); there are the

natural *homochiral* members, and then there are virtual homochiral structures obtained by artificial 'photocopying' of a structure.

4. The degree of chirality of a homochiral collection is expressible as a certain mean $S(\sigma)$ value with a variance (which at least in the case of the DLAs tested here is of normal distribution characteristics).

5. Having separate *enantiomeric populations* of right-handed, *R*, and left-handed, *S*, objects, we define an *enantiomeric match* between two individual members of these populations as $\Delta S = S(G_{improper})_R - S(G_{improper})_S$. Intuitively, one might guess that a minimal ΔS, or even $\Delta S = 0$, is indicative of nearing a situation of perfect reflectivity of both. While this may be the case, it is not necessarily so, as different fine structural details may lead to similar $S(\sigma)$ values. In fact, *isochirality*, the property of having identical $S(\sigma)$ values, does not necessarily indicate structural identity. The insensitivity of the CSM approach to detail which leads to a certain S value is a manifestation of the globality (in the thermodynamic sense) of that parameter. In this sense, S(G) can serve as a *state function*.

6. To decide on the minimal feature that dictates the chirality of the whole, one cannot adapt the approach used for small molecules. In such a case, one looks both for local centers of chirality (e.g. tetra-substituted atoms) and at the whole. In the case of the DLAs, the smallest 2D features which are chiral are four pixels arranged in the shape of an L; there is a large amount of these in a DLA, and the influence of one such L on the chirality of the whole is negligible. A larger feature must be used, and for the case of the chiral DLAs we propose to use the minimal cluster size for which the $S(\sigma)$ value begins to stabilize (at the highest resolution). For instance, in Figure 31, the minimal sizes are the chiral cores composed of $N \cong 500$ particles.

7. Let us recall Kelvin's definition of chirality: 'I call any geometrical figure, or any group of points, chiral, and say it has chirality, if its image in a plane mirror, ideally realized, cannot be brought to coincide with itself' (Kelvin, 1884).[24] Kelvin insisted in his definition that the 'plane mirror [be] *ideally* realized.' We are now in a position to lift this century-old restriction. Because we have at hand the possibility to scale chirality, an ideal mirror becomes the limiting case in the rich arsenal of natural structures.

8. Prelog's definition[23] of chirality requires that superimposability is tested by translation and rotation; we have asked above whether extension to a third type of symmetry operation, namely dilation, should also be considered. We have shown that scale invariance (dilation symmetry) of $S(\sigma)$ is one of the structural characteristics of a series of growing DLAs (Figures 31 and 32). Actually, our definition of S(G) takes care of this property, because the analyzed structure is always normalized to unit size. An important consequence of scale adjustment in the quantitative evaluation of $S(\sigma)$ is that the chirality content is an *intensive* structural property. Yet, care should be taken with this conclusion: at times one may wish to evaluate the effects of (molecular) size on physical behavior (e.g. the absorption cross section for circularly polarized light), where such scaling would be removed.

We have seen in this section that the classical definitions of chirality and its properties, which have been quite suitable for small molecules, require the introduction of more general concepts when dealing with large random objects. A generalized definition should refer to the inherent inability to form an exact counter-enantiomer; to the fact that chirality is resolution-dependent; to incidental versus inherent chirality; to the restrictive nature of an ideal mirror; to the fact that superposition can be tested by more than mere rotation-translation. Here is a tentative proposition for an extended definition, which captures many of the needs of supramolecular, mesoscopic and macroscopic structures that contain some elements of randomness in them: *Large-scale chirality is manifested by a **probe-dependent** inability to make a structure coincide with a **realization** of its mirror image by employing any proper symmetry operation. This inability is **quantifiable**.*

Three remarks conclude this section. First, we note that the chiral DLAs analyzed here may be of relevance to domains beyond chemistry. Spiral galaxies, spiral hurricane cloud formations, spiral bacterial colonies,[25] are some examples of such domains. Second, the discussion in this section, devoted to chirality which is measured through $S(\sigma)$, is immediately extendable to other chiral point groups. For instance, taking the chiral D_3 point group, which characterizes the macroscopically chiral natural quartz crystals, one can distinguish between pairs of natural enantiomers, and virtual D_3 enantiomers, obtained by reflection of a natural crystal. Third, most of the concepts discussed in this section in the context of chirality are applicable to symmetry in general. Thus, one can find in large random objects both ***incidental symmetry*** (Figure 33) and remnants of ***inherent symmetry***; and one can identify resolution dependence of symmetry in large random objects.

4. Concepts VIII - IX: Symmetry as an Order Parameter; Symmetry as a Structural Correlant with Physical Properties

Once symmetry can be quantified, the possibility exists to use it as an order parameter: the more symmetric the structure, the more ordered it is. (The link to entropy follows immediately, but we shall not comment here on continuous symmetry-entropy relations.) Furthermore, as is evident from the previous sections, symmetry can be used as a structural parameter with which one can correlate a variety of physical parameters.

We exemplify the concepts of symmetry as an order parameter and as correlant with physical parameters by summarizing a recent study[26] of cluster symmetry as a function of temperature. Considerable research has been devoted in recent years to weakly bonded clusters because of their borderline position between the condensed phases and the gas phase. In particular, cluster melting has been investigated extensively.[27] Many clusters display a melting-like transition in which the range and freedom of motion of the constituent atoms or molecules increase considerably over a relatively narrow temperature (T) range. Clusters may be trapped in the low T limit in the vicinity of a single relatively deep minimum which is often characterized by high symmetry. Then, as the temperature is raised, the cluster accesses a much

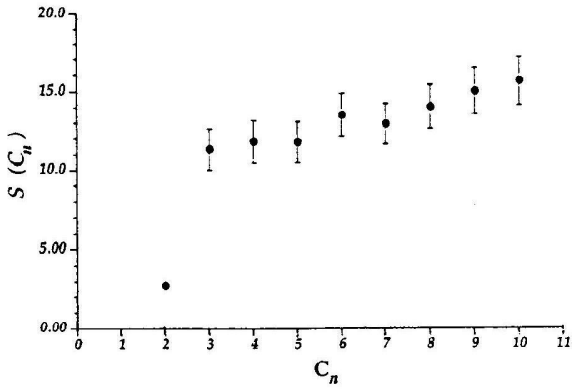


Figure 33 Incidental C_n symmetry in large random objects. Shown is the average rotational symmetry content, $S(C_n)$, with respect to $C_2 \rightarrow C_{10}$ of 158 different DLAs of the type depicted in Figure 26. It is seen that the DLAs are nearest to C_2 (the distribution bar for C_2 is small), and that the distance to all other C_n's is quite constant with a mild increase with n.

broader region of the phase space which includes a considerable number of higher-lying minima.[28]

In this context we explored the connection between cluster melting and the loss of structural symmetry of the ortho-deuterium $(D_2)_{13}$ cluster. This cluster was shown to display a melting/freezing-like transition[29] around 4K. In the low temperature limit the $(D_2)_{13}$ cluster is trapped near an icosahedral configuration, which is generally characteristic of clusters composed of 13 spherical or approximately spherical units. The cluster was simulated in the 2.5 - 5.5K T range. As T is increased, the warmed-up clusters have an increased probability to access the shallow potential wells of the very unharmonic potential energy surface. Figure 34 shows an icosahedral cluster configuration characteristic of low T, together with one of the distorted configurations obtained at 5K. (In this distorted configuration one of the molecules moved away significantly from the icosahedral position.)

The structures of the disordered clusters were first expressed with the order parameter δ. This routinely used parameter is generated by averaging over the simulation length the root mean square (rms) fluctuations in the intermolecular distances. The results were then compared to the use of symmetry as an order parameter. As a symmetry element to follow the structural changes, we selected out of the 120 elements of the icosahedral point group the inversion element, i. The reason for this selection was that as the CSM methodology always requires the minimal S(G) value, one may experience non-monotonicity in the symmetry vs physical parameter plot at a point where the S value of (say) the C_5 rotation element jumps from one nearest axis to another. This effect is minimized with inversion,

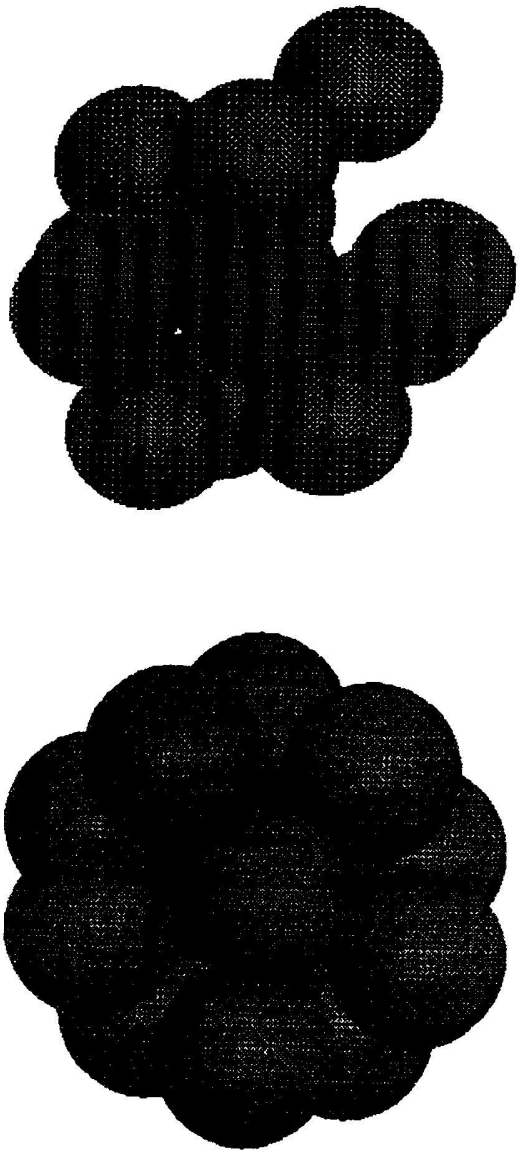


Figure 34 Bottom: the icosahedral configuration of (ortho-D_2)$_{13}$ characteristic of low temperature simulations. Top: an example of a distorted configuration obtained in a 5K simulation.

since there is only one i element per point group; hence evaluation of S(i) was selected.

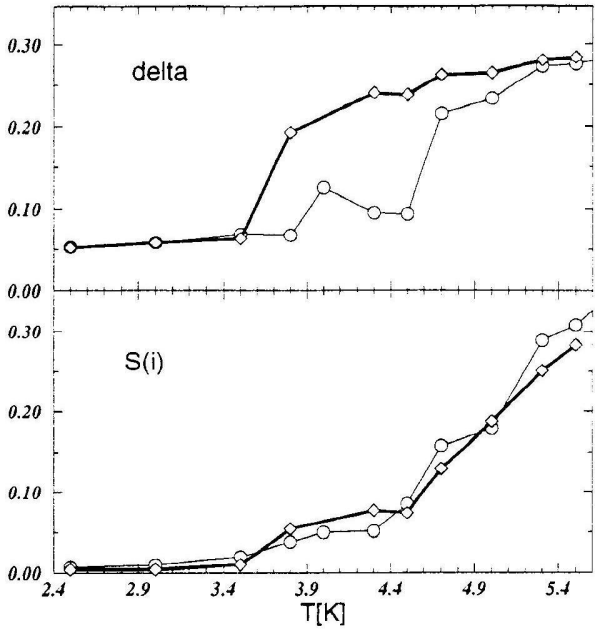


Figure 35 Top: behavior of the δ order parameter, based on rms fluctuations in the intermolecular distances. Bottom: behavior of the symmetry measure parameter S(i), which measures loss of inversion symmetry as a function of T. Diamonds: long Monte Carlo runs; circles: short runs. The robustness of S(i) vs δ is seen.

Figure 35 shows the dependence of the parameters S(i) and δ on cluster temperature. Both parameters increase around 3.4K, indicating enhanced freedom of motion suggestive of cluster melting. However the *symmetry parameter S(i) has significantly better convergence properties:* it increases gradually from zero to 0.3 in the range 3-5.5K, and the results for the two simulation lengths are nearly identical. By contrast, the results for δ (which measures the rms fluctuations in the intermolecular distances) are quite different for the two simulation lengths. In one set of results a significant jump in δ occurs at 3.7K (longer runs), and in the other at 4.6K (shorter runs). The problem with the standardly used δ is that it is expected to increase significantly at a temperature at which the atoms in the cluster start exchanging places; but one finds that the longer the simulation is, the lower the temperature at which such exchange is obtained. The S(G) parameter does not suffer from that problem. Another conceptual problem with δ is its definition, which assumes distinguishability between molecules. Distance between molecule n and n′ is not a measurable quantity in a system of N identical molecules; however, the definition of δ employs distances between *labeled* pairs of molecules. This problem is absent in the case of the S(i) parameter, because it is defined uniquely for a given

configuration of particles, and does not rely on particle labeling. Moreover, sampling the range of deviations of cluster configurations from perfect symmetry with respect to a given symmetry operation converges quite fast since we examine a range of *typical* behaviors rather than trying to sample some rare event. A limitation of the use of CSM as an order parameter is the need to select a specific symmetry operation. Thus, while all S(G) values are comparable and belong to the same scale, regardless of G, direct comparison of objects is straightforward only if a common element is selected. Nevertheless, within a defined family, say all distorted icosahedral clusters, the CSM approach seems to have interesting potential.

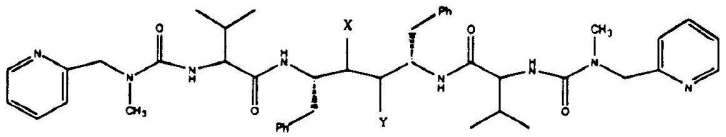
Finally we mention that S(i) proved successful in another recent study correlating symmetry with a physical property. It has been well known for some time that hyperpolarizability is governed by the deviation from centrosymmetricity. The CSM approach allowed us to link the two quantitatively for the first time. The degree of hyperpolarizability of distorted benzenes, leading to an optical non-linear effect, was correlated to the degree of centrosymmetricity of these structures.[30]

5. Concept X: Symmetry as a Structural Parameter for QSAR-Oriented Drug Design

Quantitative structure activity relations (QSAR) provides a powerful tool in the developmental studies of new drugs.[8] In these studies, attempts are made to identify predictive correlations between structural parameters of tested compounds and their activities. The concepts we consider now are symmetry and chirality, in their quantitative formulations, as investigative tools for QSAR studies.

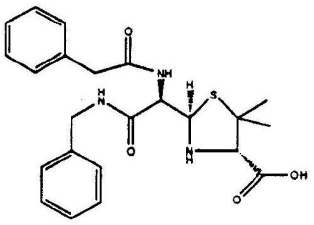
We shall treat here the degree of C_2-symmetry of anti-HIV drugs and its relation to their activity. The HIV-protease has been a major target for the development of a therapy for AIDS.[31] One significant approach has utilized the C_2 symmetry of this enzyme,[32] and in particular the C_2 symmetry of the active site. Both perfectly C_2 symmetric inhibitors, as well as inhibitors in which the C_2 symmetry was disrupted ('pseudo C_2 symmetry'), were used for that purpose. Some representative examples of studies employing this approach include those of Erickson et al.,[32] Hosur et al.[33] and Johti et al.[34] The C_2 blockers have been tailored in a way, first to optimize the fit into the active site, and second, to disrupt slightly the symmetry due to the interaction forces with the inhibitors. From these studies, it becomes evident that there is a lack of a basic measurement tool that would enable researchers in this field to treat quantitatively the C_2 structural property of the protease, its active site and the inhibitors. At the moment, researchers resort to the terms 'nearly C_2', 'almost C_2', 'pseudo C_2', 'not perfectly C_2' etc., although it is clear that, for a full picture of any QSAR, a way to measure this 'nearness' is critically needed.

Figure 36 shows examples of experimental sets of anti-HIV drugs we analyzed with the novel continuous symmetry approach. One set (A) is that of Hosur et al.[33] and the other (GR) of Johti et al.[34] The atom coordinates were taken from PDB files, and analyzed using our quantitative approach to symmetry. Series A is nearly C_2 and series GR is quite far from C_2. Some typical results of the analysis are shown in the following set of figures.

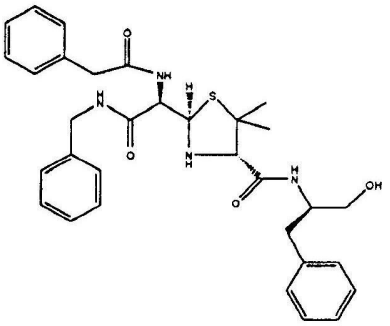


	X, Y
A-77003	R-OH, S-OH
A-76889	R-OH, R-OH
A-76928	S-OH, S-OH
A-78791	S-OH, H

GR123976



GR126045



GR137615

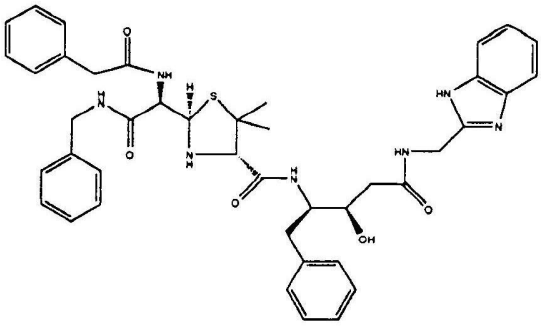


Figure 36 Two experimental sets of anti-HIV drugs. Series A is nearly C_2 and series GR is quite far from C_2.

Figure 37 shows, for the first time, a quantitative correlation between the efficiency of the HIV drug (Series A) and the C_2 distortion of the whole protein: the drug which is capable of exerting greater distortion is the more active one. Figure 38 shows a most interesting observation: the more distorted (the higher the $S(C_2)$) the drug is, the more distorted the *active site* of the protease. One affects the other. Figure 39 shows the results for the GR series. Here the trend (linear on a normal/log scale) is opposite to that of the A series, reflecting the fact that this series of drugs is so distorted from C_2 that it is quite mismatched with the C_2-ness of the active site. Therefore, in this set, the better fitting drug is the more active one.

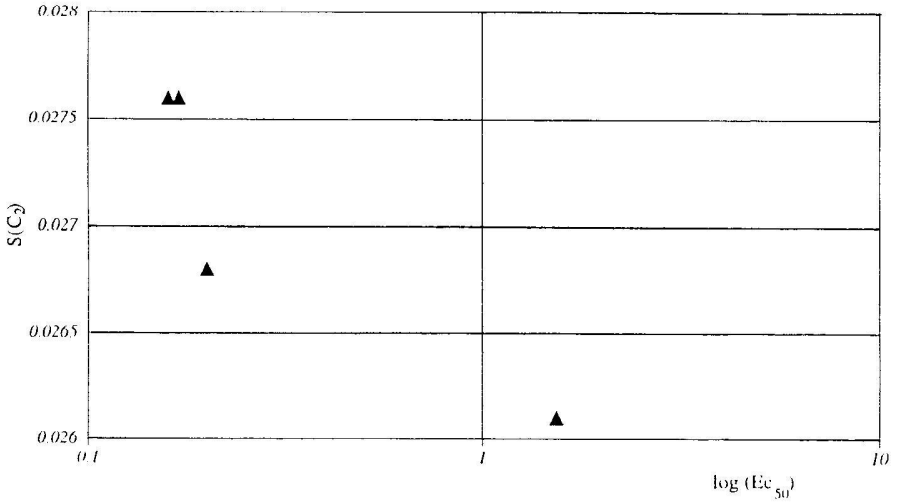


Figure 37 Shown is a correlation between the efficiency of the HIV drug (Series A) and the C_2 distortion of the whole protein.

Work aimed at developing these preliminary results into a general study of QSAR between symmetry and activity is in progress.

6. Outlook

We believe we have shown the feasibility, practicality and usefulness of enriching the descriptive language of symmetry from its primitive 'either - or' vocabulary to a vocabulary that allows shades of gray. The vast ocean that still lies ahead of us is to define the limits of the generality of this approach in identifying correlations between structure and properties, in helping us to understand the role of symmetry in such correlations, in offering useful predictions, and in phrasing new types of question regarding nature's secrets of non-symmetric structures. Such structures are the rule – not the exception.

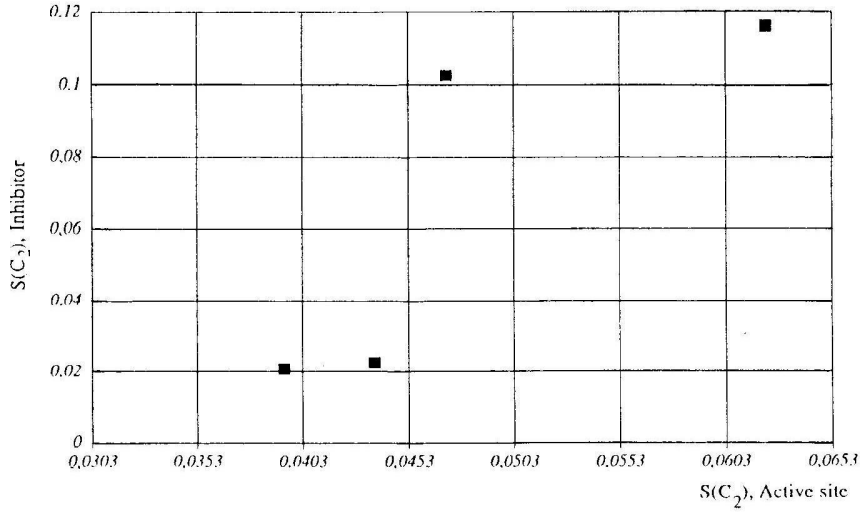


Figure 38 The more distorted (the higher the $S(C_2)$) the drug is (Series A), the more distorted the *active site* of the HIV protease.

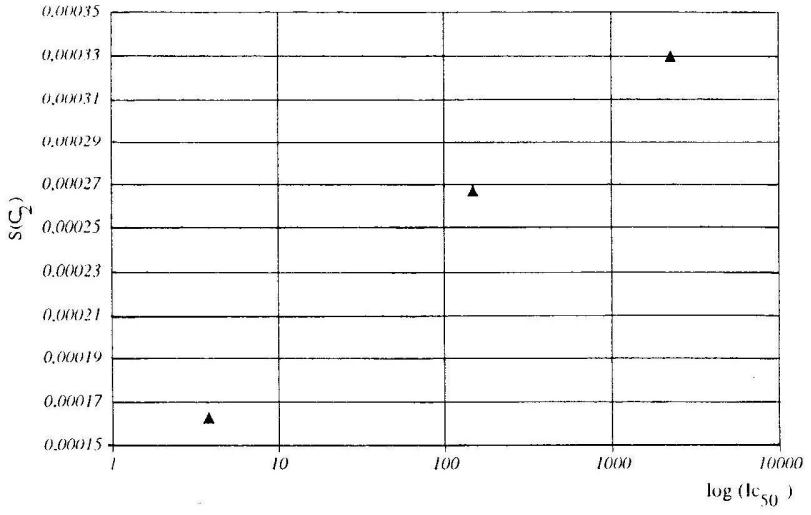


Figure 39 Symmetry/activity relation for the GR drug series.

Acknowledgments
We thank Dr V. Buch for collaboration on the use of symmetry as an order parameter and Prof. S. Shaik for illuminating discussions. The continuous symmetry project is currently funded by the Israel Science Foundation and by the Robert Szold Institute for Applied Science of the P.E.F Israel Endowment Fund. D.A. is a member of the F. Haber Research Center for Molecular Dynamics and of the Farkas Center for Light Energy Conversion.

Appendix: The Folding-Unfolding Method
The folding-unfolding concept is at the heart of the algorithm for finding the nearest symmetric shape. We demonstrate two elementary cases, which nevertheless capture the essence of the approach.

Consider first a general triangle (a scalene one), P_0, P_1, P_2 (Figure 1a).[4,6] We wish to evaluate how much C_3-ness it contains, namely how far is it from being an equilateral triangle (Figure 1c). The steps of this evaluation are based on the very steps which are used to construct an equilateral triangle, viz:

1. Determine the centroid of the triangle (Figure 40a). Translate the object so that

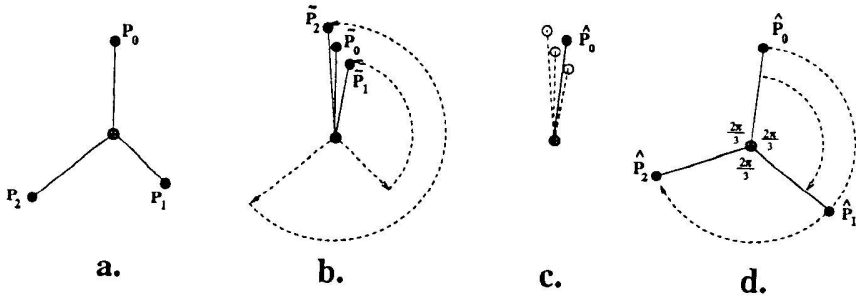


Figure 40 The folding-unfolding method explained for C_3: (a) Determine the centroid, translate the object to the origin, and scale it to size 1. (b) Fold the vertices P_o, P_1, P_2 applying to each P_i the symmetry operation g_i^{-1}. A cluster of folded points $\tilde{P}_o$, $\tilde{P}_1$, $\tilde{P}_2$ is obtained. (c) Average the folded points to obtain the average point $\hat{P}_0$. (d) Unfold the average point $\hat{P}_0$: a C_3-symmetric shape is obtained.

its centroid coincides with the origin, and scale the object to size 1 (Figure 1b).
2. Translate the symmetry group so that all operations, g_i, are about the origin (i.e. the rotations are about the origin).
3. Select an ordering of the operations of the desired symmetry group that follows the connectivity of the P_i vertices. In our case, two orderings are possible: $g_1 = E$ (the identity), $g_2 = C_3$, $g_3 = C_3^2$ and the reverse. We proceed with the first and return to the second in step 8.
4. Fold the vertices P_o, P_1, P_2, applying to each P_i the symmetry operation g_i^{-1}. A

cluster of folded points $\tilde{P}_0$, $\tilde{P}_1$, $\tilde{P}_2$ is obtained (Figures 40b, 1b). (Had the object been C_3-symmetric, all $\tilde{P}_i$'s would coincide.)

5. Average the folded points, to obtain the average point $\hat{P}_o$ (Figure 40c).

6. Unfold the average point $\hat{P}_o$ by applying to it each of the g_i operations and obtaining $\hat{P}_i$. $\hat{P}_i = g_i \hat{P}_o$ ($i = 1, \ldots, n_g$). The ordering is followed in order to retrieve the original connectivity. A C_3-symmetric shape is obtained (Figures 40d, 1c).

7. Calculate S(G) according to equation 1 (Figure 1d).

8. Minimize the S value by repeating the folding-unfolding procedure (steps 4-7) for all orderings and all orientations of the group elements. This step is equivalent to finding the best cluster of folded points. In the present case, due to the cyclic connectivity, minimization is reduced to the two orderings mentioned in step 3. As mentioned above, rigorous proof is provided[4-6] that this procedure leads to the nearest symmetric shape and results in the smallest S(G). Elsewhere we showed[6] that the procedure employed here does not translate the object, i.e. the center of mass of the set P_i, but coincides with the center of mass of $\hat{P}_i$. This means that the elements of the group are centered at the origin as well, and the nearest element sought must pass through it.

As discussed above, a natural extension of the CSM approach is to consider $S(G_{achiral})$ as a chirality measure. Our second example demonstrates how to measure chirality of a general boundary line[7,20] (Figure 41).

Given an object to be symmetrized (Figure 41a), it is converted to a necklace of an even number of boundary points, P_i, as dense as one wishes (N = 30 points in our case). When the object is already represented by vertices, this step is not needed. The treatment of the case of an odd number of points is described in ref. 7. Its center of mass is then determined and placed at the origin, and the distance from this center to the farthest P_i is scaled to 1 (Figure 41a). The aim is to find the nearest set of $\hat{P}_i$'s which is σ-symmetric, namely, to find that reflection line which will cause the set of P_i's to move minimally to the set of $\hat{P}_i$'s (Figure 41b). In the symmetrized object, each $\hat{P}_i$ must have a σ-symmetric counterpart, $\hat{P}_{N-i}$, across the reflection line (or be located on the reflection line). The full set of P_i's is divided into subsets of two points, and all possible divisions are tested. (Points on the reflection line are duplicated.) Here is how a pair of points, P_1, P_2, are σ-symmetrized with respect to a given σ (Figures 41c-41f):

1. E operates on P_1 and it remains in place ($P_1 \equiv \tilde{P}_1$); σ operates on P_2 forming the reflected $\tilde{P}_2$; a pair of adjacent points, $\tilde{P}_1$, $\tilde{P}_2$, is obtained (Figures 41c, 41d). We termed the step of applying the elements on the vertices the *folding* step. The essence of our methodology is to minimize the $\tilde{P}_1$, $\tilde{P}_2$ distance. In higher symmetry groups the cluster is much larger. For instance, in order to find the tetrahedricity measure ($S(T_d)$), a cluster of 24 points, (the number of the element in the group) is formed and minimized; see ref. 5for details.

2. The folded points, $\tilde{P}_1$, $\tilde{P}_2$ are *averaged* to $\hat{P}_1$ (Figure 41e), and $\hat{P}_1$ is

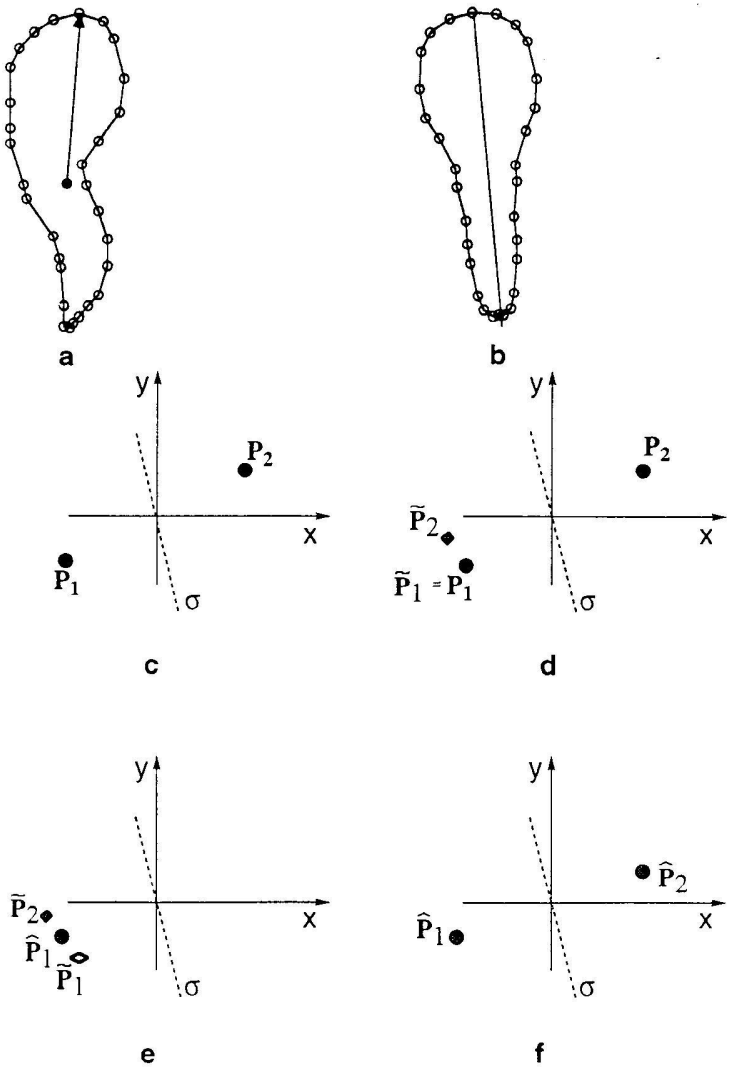


Figure 41 The folding-unfolding method explained for 2D chirality: (a) A contour composed of 30 points, P_i (the size is scaled to 1 as shown by the arrow). (b) The nearest σ-symmetric, $\hat{P}_i$, to (a). (c)-(f) Evaluation of the S(σ) value for a pair of points, P_1, P_2, with respect to a given mirror line. P_1, P_2 (c) are folded (d) to $\tilde{P}_1$, $\tilde{P}_2$, averaged (e) to $\hat{P}_1$ and unfolded (f) into the σ-symmetric pair, $\hat{P}_1$, $\hat{P}_2$.

unfolded (Figure 41f) by applying to it the elements of the group: E leaves it in place and σ forms $\hat{P}_2$ across the reflection line; the pair $\hat{P}_1$, $\hat{P}_2$ is σ-symmetric,

and the sum of distances $\|P_1 - \hat{P}_1\|^2 + \|P_2 - \hat{P}_2\|^2$ is minimal.[6] Minimization is performed by screening over all possible divisions into opposite pairs, namely over all inclinations of σ.

References

1. Gruber, B., & Iachello, L. (Eds.) 1988. 'Symmetries in Science III'. Plenum Press, New York, and earlier volumes in this series.
2. Hargittai, I. (Ed.) 1986. 'Symmetry: Unifying Human Understanding'. Pergamon Press, New York; 1989. 'Symmetry 2: Unifying Human Understanding'. Pergamon Press, Oxford.
3. Heilbronner, E., & Dunitz, J.D. 1993. 'Reflections on Symmetry in Chemistry and Elsewhere'. VCH, Basel.
4. Zabrodsky, H., Peleg, S., & Avnir, D. 1992. *J. Am. Chem. Soc.*, **114**, p. 7843.
5. Zabrodsky, H., Peleg, S., & Avnir, D. 1993. *J. Am. Chem. Soc.*, **115**, p. 8278. (Erratum: p. 11,656.)
6. Zabrodsky, H., & Avnir, D. 1995. *Adv. Molec. Struct. Res.*, **1**, p. 1; Zabrodsky, H., Peleg, S., & Avnir, D. 1995. *IEEE Trans. Pattern Anal. Machine Intelligence*, **17**, p. 1154.
7. Zabrodsky, H., & Avnir, D. 1995. *J. Am. Chem. Soc.*, **117**, p. 463.
8. Rouvray, D.H. 1995. *Topics in Current Chem.*, **173**, p. 1.
9. Readers interested in applying these programs for their own research are encouraged to contact us at david@granite.fh.huji.ac.il
10. *Bibliography of Selected Papers* on the quantitative treatment of chirality and (to a lesser extent) symmetry: Murray-Rust, P., Bürgi, H.B., & Dunitz, J.D. 1978. *Acta Cryst.*, **B34**, p. 1787; Maruani, J., & Mezey, P.G. 1987. *C.R. Hebd. Seances Acad. Sci.*, Paris, II, **305**, p. 1051 (erratum: 1988. *ibid.*, **306**, p. 1141); Mezey, P.G., & Maruani, J. 1993. *Int. J. Quantum Chem.*, **45**, p. 177; Mezey, P.G. 1989. *In* 'New Theoretical Concepts for Understanding Organic Reactions' (Eds. J. Bertran & I.G. Csizmadia). Kluwer, Dordrecht, pp. 55, 77; Gilat, G. 1989. *J. Phys. A.*, **22**, p. L545; Gilat, G. 1996. *In* 'Concepts in Chemistry' (Eds. D.H. Rouvray & E. Kirby). Research Studies Press, Somerset; Bunker, P.R. 1979. 'Molecular Symmetry and Spectroscopy'. Academic Press, New York, Chapter 11; Longuet-Higgins, H.C. 1963. *Mol. Phys.*, **6**, p. 445; Grunbaum, B. 1963. *Proc. Symp. Pre Math: Am. Math. Soc.*, **7**, p. 233; Rosen, J. 1983. 'A Symmetry Primer for Scientists'. Wiley, New York, Chapter 5; Sokolov, V.I. 1979. 'Introduction to Theoretical Stereochemistry'. Nauka, Moscow (in Russian); 1991.Gordon & Breach, Science Pubs., New York (English translation); Kitaigorodskii, A. 1961. 'Organic Chemical Crystallography'. Consultant Bureau, New York, p. 230; Weinberg, N., & Mislow, K. 1993. *J. Math. Chem.*, **14**, p. 427; Buda, A.B., & Mislow, K. 1992. *J. Am. Chem. Soc.*, **114**, p. 6006; Seri-Levy, A., & Richards, W.G. 1993. *Tetrahedron: Asymmetry*, **4**, p. 1917; Kuz'min, V.E., & Stel'mach, I.B. 1987. *Zh. Strukt. Khim.*, **8**, pp. 45, 50; Kauzmann, W., Clough, F.B., & Tobias, I. 1961. *Tetrahedron*, **13**, p. 57; Ruch, E. 1972. *Acc. Chem. Res.*, **5**, p. 49; Luzanov, A.V., & Babich, E.N. 1992. *Struct. Chem.*, **3**, p. 175; Hel-Or, Y., Peleg, S., & Avnir, D. 1990. *Langmuir*, **6**, p. 1991 (erratum: 1994. **10**, p. 1633).

11a. Mislow, K. 1954. *Science*, **120**, p. 232.

11b. Wolfe, S., Schlegel, H.B., Csizmadia, I.G., & Bernardi, F. 1975. *J. Am. Chem. Soc.*, **97**, p. 2020.

11c. Salem, L. 1971. *Acc. Chem. Res.*, **4**, p. 322.
11d. Mislow, K., & Poggi-Corradini, P. 1993. *J. Math. Chem.*, **13**, p. 209.
11e. Mezey, P.G. 1995. *J. Math. Chem.*, **17**, p. 185.
12. Pinto, Y., Zabrodsky Hel-Or, H., & Avnir, D. 1996. *J. Chem. Soc. Faraday Trans.*, in press.
13. Owicki, J.C., Shipman, L.L., & Scheraga, H.A. 1975. *J. Phys. Chem.*, **79**, p. 1794.
14. Pugliano, N., & Saykally, R.J. 1992. *Science*, **257**, p. 1937.
15. e.g. Wales, D.J. 1993. *J. Am. Chem. Soc.*, **115**, p. 11180.
16. Schütz, M., Bürgi, T., Leutwyler, S., & Bürgi, H.B. 1993. *J. Chem. Phys.*, **99**, p. 5228; Rijdt, J.G.C.M.v.D.-v., & Duijneveldt, F.B.v. 1995. *Chem. Phys. Lett.*, **237**, p. 560.
17. e.g. McConnell, H.M. 1991. *Ann. Rev. Phys. Chem.*, **42**, p. 171; Mogi, I., Okubo, S., & Nakagawa, Y. 1991. *J. Phys. Soc. Jap.*, **60**, p. 3200.
18. Meakin, P. 1992. In 'The Fractal Approach to Heterogeneous Chemistry: Surfaces, Colloids, Polymers' (Ed. D. Avnir). Wiley, Chichester, Chapter 3.1.2 (3rd corrected printing); Meakin, P. 1994. *Heterogen. Chem. Rev.*, **1**, p. 99.
19. Mislow, K., & Bickart, P., 1976/7. Isr. J. Chem., 15, p. 1.
20. Katzenelson, O., Zabrodsky Hel-Or, H., & Avnir, D. 1996. *Chem. Europ. J.* (in *Angew. Chem. Int. Ed. Eng.*), **2**, p. 174.
21. Schwartz, D.K. 1993. *Nature*, **362**, p. 593.
22. Mezey, P.G. 1993. 'Shape in Chemistry'. VCH, Weinheim.
23. Prelog, V. 1975/6. J. Molec. Catal., 1, p. 159.
24. Lord Kelvin. 1884. 'Baltimore Lectures'. p. 436 (c.f. also 1904. 'Baltimore Lectures'. Appendix H, p. 439; Whyte, L.L., 1958. *Nature*, **182**, p. 198.)
25. Ben-Jacob, E., Shocet, O., Tenenbaum, A., Cohen, I., Czirok, A., & Vicsek, T. 1994. *Fractals*, **2**, p. 15.
26. Buch, V., Greshgoren, E., Zabrodsky Hel-Or, H., & Avnir, D. 1995. *Chem. Phys. Lett.*, **247**, p. 149.
27. e.g. Buch, V. 1994. *J. Chem. Phys.*, **100**, p. 7610, and references cited therein.
28. Beck, T.L., & Berry, R.S. 1988. *J. Chem. Phys.*, **88**, p. 3910.
29. Scharf, D., Martyna, G.J., & Klein, M.L. 1992. *Chem. Phys. Lett.*, **197**, p. 231.
30. Kanis, D.R., Ong, J.S., Marks, T.J., Ratner, M.A., Zabrodsky, Z., Keinan, S., & Avnir, D. 1995. *J. Phys. Chem.*, **99**, p. 11061.
31. Kim, E.E., et al. 1995. *J. Am. Chem. Soc.*, **117**, p. 1181.
32. Erickson, J., et al. 1990. *Science*, **249**, p. 527.
33. Hosur,M.V., et al. 1994. *J. Am. Chem. Soc.*, **116**, p. 847.
34. Jhoti, H., et al. 1994. *Biochemistry*, **33**, p. 8417.