Effects of temperature on the shape and symmetry of molecules and solids

Abel Carreras[a], Efrem Bernuz[b], Xavier Marugan[b], Miquel Llunell[b], and Pere Alemany [b,]*

[a] Donostia International Physics Center (DIPC), Paseo Manuel de Lardizabal 4, 20018 Donostia, Euskadi, Spain

[b] Departament de Ciència dels Materials i Química Física and Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, Diagonal 647, 08028 Barcelona, Catalunya, Spain.

* Author to whom correspondence should be addressed.

Tel: +34 93 403 1239; fax: +34 93 402 1231; e-mail: p.alemany@ub.edu

Abstract

Despite its undeniable problems from a philosophical point of view, the concept of molecular structure, with attributes such as shape and symmetry, directly borrowed from the description of macroscopic objects, is nowadays central to most of the chemical sciences. Descriptions such as "the tetrahedral" carbon atom or "octahedral coordination complexes" are widely used as much in elementary textbooks as in the most up-to-date research articles. The definition of molecular shape is, however, not as simple as it might seem at first sight. Molecules don't behave as macroscopic objects do, and the arrangement of atoms within a molecule changes continuously due to the incessant motion of its constituent particles, nuclei and electrons. How are molecular shape and symmetry affected by this thermal motion? In this review we introduce the language of continuous symmetry measures as a new tool to quantitatively describe the effects of temperature on molecular shape and symmetry, enriching in this way the set of molecular descriptors that might be used in the establishment of new empirical structure-property relations, of great interest in concomitant areas such as medicinal chemistry or materials science.

The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.

Sir Lawrence Bragg

1 Introduction

Molecules depicted in modern chemistry textbooks have distinct physical sizes, shapes, and structures. They are composed of individual atoms linked by bonds and usually represented by mechanical models such as ball-and-sticks or space-filling CPK models, or in more recent times by computer rendered interactive 3D images of these classical models (Figure 1 left and centre). As noted by Ramberg[1], the construction and use of these representational models is of central importance in the development of chemical theories since they make visible the invisible world of atoms, providing it with a graphic clarity or *Anschaulichkeit*. Techniques for modeling molecules and atoms date at least to Dalton's wooden spheres, although these early models were originally not meant to visualize molecules in a physical sense, ignoring in some cases the representation of the atom itself as in van't Hoff's stereoformulae (Figure 1 right), where the arrangement of atoms around a carbon atom was represented by a tetrahedron, but neither the central atom itself nor the bonds between this atom and the atoms in its coordination sphere were explicitly shown.

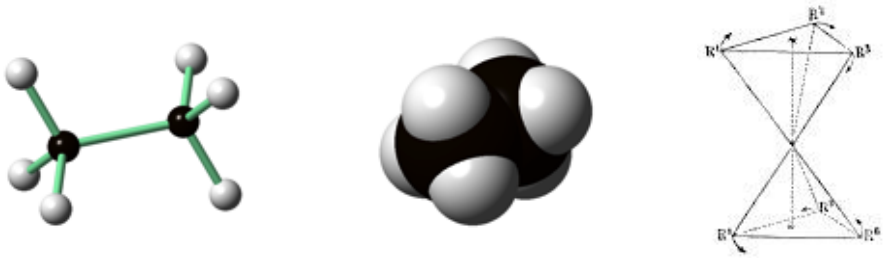


Figure 1. From left to right: ball-and-stick model, space filling model, and van't Hoff's stereoformula for ethane. Note that in van't Hoff's stereoformulae the ligands around two carbon atoms linked through a simple C-C bond were usually represented in the eclipsed rather than in the staggered conformation.

It was short after the introduction of the theory of the tetrahedral atom by van't Hoff in 1874[2] that chemists began to consider those models as a true physical representation of the microscopic world. This transformation in meaning, largely completed by the turn of the twentieth century, gave birth to stereochemistry, a new sub-discipline of chemistry concerned with molecules as entities with shape and structure. The classical concept of molecular structure underlying stereochemistry has, however, been questioned adducing that it has no physical reality at all since it cannot be derived directly from the physical laws (quantum mechanics) governing the motions of the nuclei and electrons.[3] Its utility in reasoning about microscopic events has been, nevertheless, at the base of its undeniable success in chemistry.[4] Mechanical concepts such as that of the "steric hindrance" of bulky groups allow chemists to discuss reaction mechanisms in vivid terms and to convince us that we are actually witnessing what is going on in a reaction flask at the microscopic level.

Shape and symmetry of molecules, crystals, supramolecular aggregates or local coordination environments around individual atoms are the essential concepts used in stereochemical analyses, where these shapes and symmetries are described using the tools of classical Euclidean geometry and group theory. Since the beginnings of stereochemistry, polyhedra, especially high symmetry convex polyhedra such as the Platonic solids, have played a key role in the rationalization of stereochemical information[5] and nowadays concepts such as the "tetrahedral carbon atom" or "octahedral coordination complexes" form part of the common knowledge of every chemist.

The success of these simple models in explaining the properties and reactivity of large families of compounds has, however, the undesired side effect that most chemists assume naturally that these ideal symmetrical shapes are found for the vast majority of molecules, with low-symmetry cases being considered as unavoidable exceptions due to the complexity of nature. A careful analysis of experimentally determined molecular structures, greatly facilitated by today's structural databases[6] shows, that the situation is exactly the opposite: the presence of ideal symmetrical shapes is an exception, rather than the general rule. This situation was already described by Baur in 1974[7] in a careful examination of over 200 well-resolved crystal structures, both of organic and inorganic compounds, containing the presumably tetrahedral phosphate anion. In his study Baur was unable to find a single case with perfect tetrahedral symmetry for the PO_4^{3-} group, and what is even more striking, in almost 85% of the analyzed crystals, the phosphate anion did not show any symmetry at all. In just 2% of the structures one of the four C_3 rotation axes of the tetrahedron was still present, and the highest symmetry, D_{2d}, was found for a single case.

This discrepancy between idealized and real molecular structures is of course inherent to the very nature of molecules and to the representation of their structures in terms of objects, the molecular models, for which the notion of shape usual in our everyday environment is valid. Molecules are, however made of an ensemble of particles, nuclei and electrons, dancing around each other in a continuous movement. This dynamic nature is very distant to that of the solid objects in our macroscopic world for which the mathematical notions of shape and symmetry were developed, and a rigorous stereochemical analysis should take this fact into account. The basic sources of departure of real molecular structures from their idealized shapes are basically two: many molecular structures are easily deformable by either the presence of other molecules in their close environment, or simply by their dynamical nature and the effect of temperature that results in a constantly changing structure amenable only to experimental scrutiny via time or space averaged structures that are often more symmetric than the instantaneous structures between which the molecule is constantly fluctuating. In this review we will address this later problem, the effect of temperature on molecular structures, showing how the combination of molecular simulation techniques and the description of shape and symmetry as continuous properties allows us to gain a deeper knowledge of the concept of molecular structure, introducing some nuances that are often overlooked in traditional stereochemical studies. Since a vast literature (the most relevant references are given in the following sections) on using this approach to reveal the effects of the crystal environment on

the molecular structure is already available, we will focus our attention here exclusively on the effects of temperature on molecular shape and symmetry.

2 Molecular shape and symmetry

Before analyzing the effects of temperature on molecular shape and symmetry we should clearly establish what we mean exactly by "molecular shape". In common language, the word “shape” is attached to a personal intuition of what each of us thinks that defines a macroscopic object in three-dimensional (3D) space. In some cases, for instance with relatively simple human made objects such as a buildings, we make an abstraction and we may view the object as a distribution of points in space describing its shape in terms of distances and angles, that is, local geometrical measures. More often, for instance when we try to describe the shape of natural objects such as animals, we associate the notion of shape with an envelope surface. In this case, rather than a precise geometrical characterization, it may be sufficient to describe shape in more qualitative terms or using comparisons with well known objects, for instance when we say that somebody has a pear-shaped figure. Both types of "shapes", represented by classical ball-and-stick and space-filling models respectively (Figures 1a and 1b), are used in chemistry, depending on which molecular properties one chooses to emphasize. Ball-and-stick models are better suited to represent the spatial arrangement of atoms within a molecule, while space-filling models give a more accurate description of the excluded volume occupied by a molecule and are better suited to study questions such steric hindrance or the fitting of a substrate in an enzime's cavity.

Since in this review we are mostly interested in the effect of temperature on geometrical aspects of molecular structure, we will use basically the first type of representation of molecular shape, a ball-and-stick model. From a mathematical point of view our molecules will, thus, be represented by a collection of points in Euclidean 3D space corresponding to the fixed position of the atomic nuclei in a given molecular conformation. In some cases we might be interested in distinguishing the points in this set according to the different nuclei they represent, in other cases we might just be interested in comparing shapes between different molecules with related structures and we do not make distinctions between the nature of the atoms forming these molecules. This representation of the molecular structure is very convenient for the stereochemical analysis of experimental structures determined by x-ray or neutron diffraction techniques, which are readily accessible from databases such as the Cambridge Structural Database (CSD),[6a] the Protein Data Bank (PDB),[6c] or the Inorganic Crystal Structure Database (ICSD).[6b]

From the point of view of theoretical chemistry it is customary to consider that a given molecular structure can be directly inferred from quantum chemical calculations just by obtaining the set of nuclear coordinates that leads to a minimum on the potential energy surface of the system, although it has been emphasized[3b] that physical systems are never isolated nor closed, and for this reason, as with the size of individual atoms, so the geometry of a molecule varies depending on their environment. Today we know that despite this inherent difficulty, the use of these geometries to discuss chemical problems makes sense if we do not forget that they are relative to the timescale of measurement. In this sense, the notion

of molecular structure based on “frozen” nuclear conformations is limited because from quantum mechanics we know that the vibration of nuclei around their minimal energy position cannot be eliminated even at the lowest attainable temperatures, and consequently, an accurate description of molecular shape should include a “smearing effect” resulting from nuclear motion. Different solutions have been proposed in the literature to address this question.[8] Basically two different approaches are available. One is to use dynamical shape descriptors, which take into account the dynamical nature of molecular shape due to the nuclear flexibility. Alternatively one may use static shape descriptors obtained for a frozen nuclear geometry and use molecular simulation techniques to generate a representative series of frozen nuclear geometries to study the dependence of these shape descriptors with external factors such as temperature in a statistical way. In this review we adopt this second point of view, by using the language of continuous shape and symmetry measures defined in the next section to characterize the shape and symmetry for a set of structures of a given molecule obtained by either Monte Carlo or molecular dynamics simulations at a constant temperature.[9]

3 Continuous shape and symmetry measures

We all have a notion of shape, which is already acquired in our preschool years since it is generally accepted that understanding color and shape is a basic tool for learning many skills in all curriculum areas, from math and science to language and reading. Associated with the concept of shape we also learn very early to recognize symmetry: the bilateral symmetry in our bodies, the rotational symmetry of many flowers, or the complex, but at the same time apparently simple symmetry of polyhedral shapes such as the cube. It should be thus no surprise that polyhedra and symmetry have entered as a basic ingredient in many of our conceptual models developed to explain the physical world. The extensive use of geometry and symmetry in modern science are a consequence of the deep influence of Greek philosophy in western culture, starting with Plato’s attempt to relate the properties of matter with those of the highly symmetrical regular solids. The modern use of shape and symmetry in scientific theories derives from the Erlangen program proposed by F. Klein in 1872 to use group theory, a branch of mathematics that relies on algebraic methods to abstract the idea of symmetry, as the most useful way of organizing geometrical knowledge.[10] According to this philosophy, the shape (symmetry) of an object, for instance a molecular structure **Q** formed by a set of points in Euclidean space, is best described by those transformations (automorphisms) of space which leave **Q** invariant. One can show that the set of these automorphisms has the structure of a mathematical group, called the object's symmetry group *G*. It is easy to convince ourselves just by playing with a model of a cube that the shape (symmetry) of the cube is preserved by translations, rotations, and scaling transformations. Although its appearance to our eyes might change, a cube is a cube independently of its position, its orientation, and its size, and we are able to recognize it as a cube independently of our viewpoint. More subtle transformations preserving the identity of a cube are reflections, but from our particular experience we all know that we are also able to recognize the shape of a cube just by looking at its image on a mirror.

At this point it is important to note a subtle, but important, difference between shape and symmetry. For regular polyhedra such as the tetrahedron, shape and symmetry are univocally linked since the

tetrahedron is the only possible 4-vertex polyhedron in Euclidean 3D space having tetrahedral symmetry. This is no longer true if we consider, for example, a rectangular prism. Just by changing the relative lengths of the three different edges we may have many particular shapes, all with the same symmetry. In this respect, shape is a more restrictive concept than symmetry: not all objects with the same symmetry have necessarily the same shape.

Let us leave for a moment the abstract mathematical description of shapes and symmetries and return to stereochemistry. Much of the actual stereochemical wisdom relies on old venerable hypotheses such as van't Hoff's and Le Bel's postulate that the four valencies of a carbon atom are arranged in a tetrahedral disposition. From the point of view of a structural chemist in the 21st century, with access to the wealth of structural information contained in data bases, it should be, however, remarkable that such theories, derived for idealized symmetric molecular geometries before it was even possible to determine the molecular structure, can be applied without major revisions to obtain rational explanations for many experimental observations, even if, as noted above, we have evidence that the structures of most molecules are not symmetric at all.

It is precisely when trying to understand this apparent contradiction that one of the major shortcomings of the group theoretical approach to symmetry becomes evident. If one associates symmetry just with the group of spatial transformations that leave an object invariant, symmetry and shape become automatically dichotomous properties, that is, a molecule has either a certain symmetry (shape) or not. In this respect, even tiny atomic displacements, which might have an imperceptible influence on the physical properties of the molecule, are sufficient to destroy its symmetry (shape). A satisfactory solution for this problem can only be given if the mathematical treatment of symmetry and shape is radically changed by introducing a metric notion, the degree of a given symmetry or shape present in an object, that allows to overcome the usual 'black or white' description of symmetry in group theory by introducing a continuous scale containing all infinite shades of gray between these two extremes. Although different approaches to this problem are possible[11] we will use here the so-called continuous symmetry or shape measures (CSMs or CShMs) introduced by Avnir and coworkers in the early 1990s.[12] [13] CSMs and CShMs are especially well suited for its use in chemistry because, although originally devised to describe the geometric symmetry of molecular structures, they can be easily generalized[14] to treat the degree of symmetry of more complex mathematical objects commonly used in quantum chemistry[15] such as wave functions,[16] orbitals,[17] [18] [19] electron densities,[20] or even quantum mechanical operators.[21] We will, however in this review, limit ourselves to give a brief introduction to the mathematical foundations of geometrical CSMs and CShMs and use them afterwards to show how they can be used to address, in a quantitative manner, the effects of temperature on molecular or solid state structures. The reader interested in a detailed application of CShMs and CSMs to stereochemical problems will find more information in the extensive literature included in the reference list. The main questions addressed with the CSMs formalism include the precise description of the coordination geometry of individual atoms in either molecules [22] [23] [24] [25] [26] or solids,[27] the symmetry of clusters,[28] supramolecular assemblies,[29] or proteins,[30] chirality in transition metal compounds,[31] [26b] aggregates,[32] or in solids,[33] and the relation

between symmetry and the physical properties of molecules[34] and solids.[35] CMSs have also been used to describe geometrical changes along reaction paths,[36] particularly for enantiomerization processes.[37] An interesting application of CSMs, closely related to the type of work described here is the analysis of environmental effects on molecular symmetry,[38] with a special emphasis on the coexistence of apparently contradictory symmetries in a same system.[27d, 39]

Since a detailed description of the algorithms used for computing CShMs and CSMs is readily available,[40] we will give here only a brief overview of the main points, focusing more on the ideas behind the CSM formalism than on mathematical technicalities. The most intuitive approach to understand the mathematical formalism of CShMs is to think on how we would proceed if we were presented with an object, for instance a quadrangle **Q** as in figure 2, and we were asked about how much it resembles another object $\mathbf{P_0}$ with a given ideal shape such as a square. Given that the shape is invariant upon translations, rotations, and scaling, the most evident way to compare the two objects is to translate, rotate and scale one of them, for instance $\mathbf{P_0}$, until we maximize the overlap between **Q** and **P**, the image of $\mathbf{P_0}$ after these transformations (Figure 2, right). Once we have determined **P** we can define the continuous shape measure of **Q** with respect to the ideal shape **P** as:

$$S_P(Q) = \min \frac{\sum_{i=1}^{N} |q_i - p_i|^2}{\sum_{i=1}^{N} |q_i - q_0|^2} \times 100 \qquad [1]$$

where N is the number of vertices in the structures we are comparing, q_i and p_i are the position vectors of the vertices of **Q** and **P**, respectively, and q_0 the geometric center of the problem structure **Q**. The minimization in eq. [1] refers to the relative position, orientation, and scaling that must be applied to $\mathbf{P_0}$ to maximize the overlap (or alternatively, to minimize the squares of distances between their respective vertices, which is actually the case in eq. [1]). If the matching of the two shapes is described, as in eq. [1], by the distance between vertices of the two objects, a further minimization with respect to all possible ways to label the N vertices in the reference structure $\mathbf{P_0}$ is also needed. While an analytical solution to find the best relative position, orientation and relative size of the two objects is known, the permutation problem must be solved by sweeping through all possible $N!$ permutations of labels if all vertices are equivalent, or to restricted sets of $N_1! \times N_2! \times \ldots \times N_k!$ permutations if the set of N vertices can be partitioned into k sets with $N_1, N_2, \ldots, N_k$ equivalent vertices each.

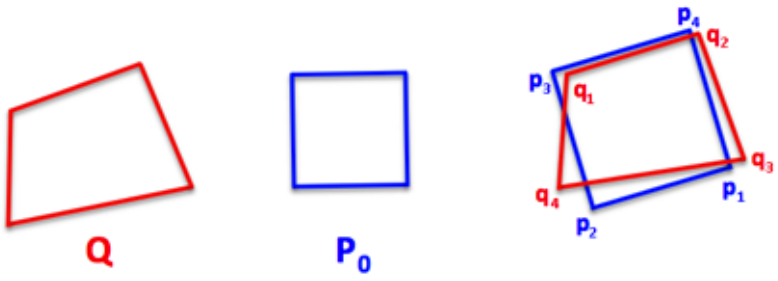


Figure 2. Procedure used to calculate the CShM of an arbitrary quadrangle **Q** with respect to the square $\mathbf{P_0}$ by minimizing the squares of the distances between the vertices of the two structures with respect to translations, rotations, scaling and vertex labeling.

From eq. [1] it follows that if **Q** and $\mathbf{P_0}$ have exactly the same shape, then $S_P(Q) = 0$. Since $S_P(Q)$ is always positive, the larger its value, the less similar is **Q** to the ideal shape $\mathbf{P_0}$. It can be shown that the maximum value for $S_P(Q)$ is 100, corresponding to the unphysical situation where all vertices of **Q** collapse into a single point. Although, in principle, CShM values can fall in the range between 0 and 100, measures for severely distorted chemical structures are never larger than 50 and as a rule of thumb we can consider that chemically significant distortions give shape or symmetry measures about 0.1 or higher, while values larger than 3 indicate important distortions. For structures with CSMs values above 5 the use of the ideal shape $\mathbf{P_0}$ to describe the shape of a given object is already strongly questionable.

It is important to notice that describing the shape of a geometrical object using a CShM represents a very strong reduction of information, from the *3N-6* internal coordinates needed to fully specify the relative positions of its vertices to just a single shape descriptor. For small distortions the use of a single CShM, the one corresponding to the closest ideal structure, is a valid simplification, but if the structure is severely distorted, far from any simple highly symmetric structure, the best description for the shape needs the specification of several CShMs with respect to different reference shapes. A special case appearing often in stereochemistry is that of a structure that falls along the interconversion path between two high symmetry structures.[41] Let's consider, for instance, the so-called "spread" pathway between a tetrahedron and a square (Figure 3a) obtained by flattening the tetrahedron while maintaining a D_{2d} symmetry for the structure. Structures along this path have tetrahedral CShMs comprised between 0 and 33.3. The shape of molecular structures with small values of $S_{Td}(Q)$ are well described by the tetrahedral measure, but structures close to the square planar one with $S_{Td}(Q) \approx 33$ obviously not, since this high value just indicates that we are far away from the tetrahedron, but not in which direction. To describe these structures we will better use the square planar shape measure $S_{Sq}(Q) \approx 0$. For structures in the middle part of the path it will not be possible to use a single high symmetry reference structure and we will need to indicate, at least, both $S_{Td}(Q)$ and $S_{Sq}(Q)$. A very practical way to present this information visually is the use of shape maps[25b] such as the one shown in figure 3b where the value of a CShM $S_P(Q)$ with respect to ideal shape **P** is represented as a function of $S_R(Q)$, the shape measure for the same structure with respect to another ideal shape **R**. The use of shape maps and the so-called minimal distortion pathways between two ideal structures[41] such as the "spread" path between the tetrahedron and the square, have been shown to be of great utility in the rationalization of large quantities of stereochemical information, especially in the description of the coordination geometry around metal atoms in complexes and solids.[22a, 23c, 24-25, 25f, h]

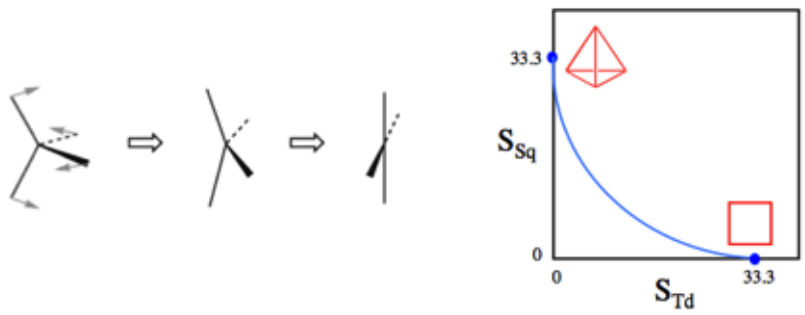


Figure 3. a) Spread distortion for a tetrahedral ML_4 structure. b) Tetrahedral-square planar shape map showing the location of the perfect tetrahedral and perfect square geometries and the minimal distortion interconversion path (the spread distortion) between them.

To define a continuous scale to gauge the degree of symmetry of an object described by a set of vertices we may proceed in the same way as for the definition of CShMs. The final result for the symmetry measure with respect to a given point symmetry group G, denoted as $S_G(Q)$, is an equation totally analogous to eq. [1], where **Q** refers again to our problem structure but where **P** is now the G-symmetric structure closest to **Q**. The minimization process in this case refers to the relative position of the two structures (translation), the orientation of the symmetry elements for the reference G-symmetric structure **P**, the scale factor, and again the labeling of vertices of the symmetric structure. Note that although the same equation may be used both to define shape and symmetry measures, there is a fundamental difference between the two procedures: while in computing a shape measure we know in advance the ideal object $\mathbf{P_0}$, for instance a tetrahedron if we want to compute $S_{Td}(Q)$, in the case of symmetry measures the shape of the closest G-symmetric structure is, in principle, previously unknown. Consider, for instance that we would like to measure the rectangular symmetry for a given general quadrangle. Besides optimizing to seek for the translation, rotation and scaling that leads to the optimal overlap of our quadrangle **Q** with a given rectangle, we will need to consider also which is the best rectangle by optimizing the ratio between the lengths of the two different edges of the rectangle. Although this additional optimization process may seem difficult to generalize for any given symmetry group, it has been shown that it is possible to do it efficiently using either the folding-unfolding algorithm[13a] or via the calculation of intermediate symmetry operation measures.[40a, b] As in the case of shape measures, the values of CSMs are also limited to the $0 \leq S_G(Q) \leq 100$ interval with $S_G(Q) = 0$, meaning that Q is a G-symmetric shape. In some special cases as those of the platonic solids described below, it can be shown that symmetry and shape measures are equivalent, a consequence of the fact that a given regular polyhedron is the only possible polyhedron with a given number of vertices and a given symmetry. The octahedron, for example, is the only possible six-vertex polyhedron with octahedral symmetry. This is, however, not true for most of the polyhedra, for which shape and symmetry are not univocally linked.

A special mention should be made to chirality,[12c] a specific type of symmetry that has a prominent role in chemistry. A chiral object is usually described as an object that cannot be superposed with its mirror image. Technically speaking, chirality implies lack of improper rotation symmetry and its CSM can be based on estimating how close a chiral object is from having this symmetry. Using the CSMs defined above, the continuous chirality measure (CCM) can be defined as the minimal of all S_G values for $G = S_n$

with n = 1,2,4.... In most cases it will be either $S_{chir} = S_{Cs} = S_{S1}$ or $S_{chir} = S_{Ci} = S_{S2}$, while in a few cases we will have to look for S_{S4} or higher-order even improper rotation axes. Since in most cases visual inspection of the studied structure is enough in order to guess which one could be the nearest S_n group, a practical solution is just to calculate this particular S_{Sn}, or in case of doubt a few S_{Sn} values and pick the smallest one.

A chiral molecule may exist in two different structures that are one the mirror image of the other. These two structures, called enantiomers, have, of course, both the same chirality measure, that is, they are both at the same distance from the closest achiral structure and have the same shape, which means that they are both equidistant from a given reference shape. To distinguish both enantiomers we must introduce the notion of handedness, which should not be confused with chirality itself.[42] A given chiral structure may be labeled as right- or left-handed, independently of the degree of chirality that it possesses. We should however recall that labeling an object as left- or right-handed is an arbitrary convention. In chemistry different standard methods to assign handedness are being used for different cases. The most well known are the Cahn–Ingold–Prelog (CIP) priority rules,[43] the rules for chiral octahedral six-coordinated complexes,[44] those for tris-bidentate octahedral complexes,[45] or the rules for helical chirality which require the existence and identification of a C_2 symmetry axis.[46] When one approaches the handedness-labeling of a family of objects which is not covered by existing labeling schemes, due to the arbitrariness of the problem, all options are open. In this review we will describe the effect of temperature on the local atomic environment for tetrahedral solids such as diamond. Since a distortion from a regular AB_4 tetrahedron leads, in general, to a chiral structure, we will use the rule devised by Avnir and coworkers to assign a left/right handedness to AB_4 chiral tetrahedra. The interested reader is referred to the original reference[42] for more details on this aspect.

4 Effect of temperature on shape and symmetry

We will discuss the principal effects of temperature on molecular shape and symmetry by means of a simple example, the highly symmetric tetrahedral P_4 molecule. Optimization of the molecular structure using the semiempirical PM6 method[47] yields a perfectly tetrahedral geometry with a P-P distance of 2.26Å. This geometry is confirmed to correspond to a minimum on the potential energy surface for which six normal modes with 938 (A_1), 626 (T_2), and 431 cm^{-1} (E) wavenumbers are found in a normal mode analysis. As mentioned earlier, temperature will cause local vibrational motion of the atoms and at any time the P_4 molecule will adopt a distorted tetrahedral structure, in most cases with no symmetry at all. We can make an estimation of the size of the local atomic motions by using the analogy of the vibrations of a mass *m* held in place by springs.[48] The energy associated with its displacement is $m\omega^2\langle u^2\rangle$ where *u* is its displacement from the equilibrium position and ω the angular frequency of the oscillatory motion. Equating this mean displacement energy with the thermal energy k_BT gives:

$$\left\langle u^2 \right\rangle^{1/2} = \sqrt{\frac{k_B T}{m\omega^2}} \qquad [2]$$

For the P_4 molecule at T =300K, if we consider the intermediate normal mode T_2 with $\omega/2\pi$ = 18.8 THz we obtain a value of 0.024Å for $\langle u^2 \rangle^{1/2}$, that is, about 1% of the P-P distance in the equilibrium geometry. Considering a single P-P bond, its distance will oscillate between 2.21 and 2.31Å and the 60º angle spanned by three P atoms on a face of the tetrahedron will suffer changes of about $\pm$0.6º. Are these changes important for the loss of tetrahedral symmetry? As we will see below, the average tetrahedral symmetry measure at 300 K is about 0.06, indicating that the average distortion is small, but non negligible.

Before analyzing in detail the evolution of the symmetry loss with temperature, let us note three conclusions that are readily available from this qualitative analysis. First of all, from equation [2] we can deduce that the average atomic displacement will be larger for light than for heavier atoms, and it will also be larger in structures with low energy (low frequency) vibrational modes. In this respect, for a given structure in general we find that twisting modes will cause larger symmetry losses than bending ones, while high frequency stretching modes will lead to the smallest distortions. The second noteworthy point is that since shape and symmetry are invariant under scaling, the important parameter to gauge the size of a distortion is not the absolute value of the displacement, but the relative size of this displacement with respect to the average size of the structure. In other words, a 0.02Å average atomic displacement in a P_4 tetrahedron with 2.26Å edges will lead to smaller symmetry losses than the same displacement in a tetrahedral C_4 structure with 1.51Å edges. The combination of these two effects, the displacement dependence on mass and frequency on one hand and the size of the structure on the other hand, will be shown to play an important role in the unexpected behavior found for the symmetry loss in tetrahedral solids such as carbon, silicon and germanium. The third important point that should not be overlooked is that we are discussing about the average shape (symmetry) loss that a structure will suffer due to thermal motion. This magnitude behaves, however in a radically different way from the shape (symmetry) of the average structure. Considering these two magnitudes leads to a striking contradiction often found in crystallography: raising the temperature, and hence, enlarging the displacement of individual atoms from their equilibrium positions, leads, in general, to more symmetric structures. This situation is nicely illustrated by comparing the crystal structures at different temperatures for alkali-metal cyanides or alkaline-earth carbides such as NaCN[49] or CaC_2.[50] While these compounds have low temperature orthorhombic, tetragonal, or monoclinic structures where the diatomic CN^- or C_2^{2-} dumbbells are all aligned in the same direction, raising of the temperature allows a rotation of these units and the observed high temperature structures become cubic salt-rock type structures where each diatomic anion occupies a single position in the crystal corresponding to the anion's mass center. Thermal motion leads in this case to a more symmetric average structure, the instantaneous symmetry of the structure at any given time is however lower. In this respect, the symmetry enhancement produced by temperature is a consequence of the limitations of our observation tool, x-ray analysis, which allows us to explore the structure of a crystal in a time scale that is too slow to appreciate the instantaneous symmetry breaking taking place at the microscopic scale. In our simulations we also found that despite having an average tetrahedral symmetry measure of 0.06 at 300K, the average structure for the P_4 molecule is perfectly tetrahedral within numerical error. In fact this result is a confirmation of the goodness of the harmonic

approximation that is used to describe atomic motion within a molecule in terms of normal vibration modes. If all atoms are subject to perfectly harmonic motions around their equilibrium positions, the average structure will, per force, maintain the shape, symmetry, and even the size of the equilibrium structure. In our simulations we find that while the shape and the symmetry of the average structure are maintained as the temperature is raised, its size increases slightly due to anharmonicity.

Let us now analyze in more detail how the perfect tetrahedral symmetry of the P_4 molecule is lost as we consider the motion of each P atom around its equilibrium position.[51] Figures 4a and 4b show the distribution of total energies and tetrahedral CSM for the set of 30.000 P_4 structures at a temperature of 300 K, respectively. As it has been described earlier by Tuvi-Arad and coworkers,[36f, 52] a fit of the histogram to an analytic function leads to a log-normal distribution, eq. [3], both for the energy and the CSM,

$$P(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \qquad [3]$$

where μ and σ are two parameters related to both the position of the peak and the width of the distribution.

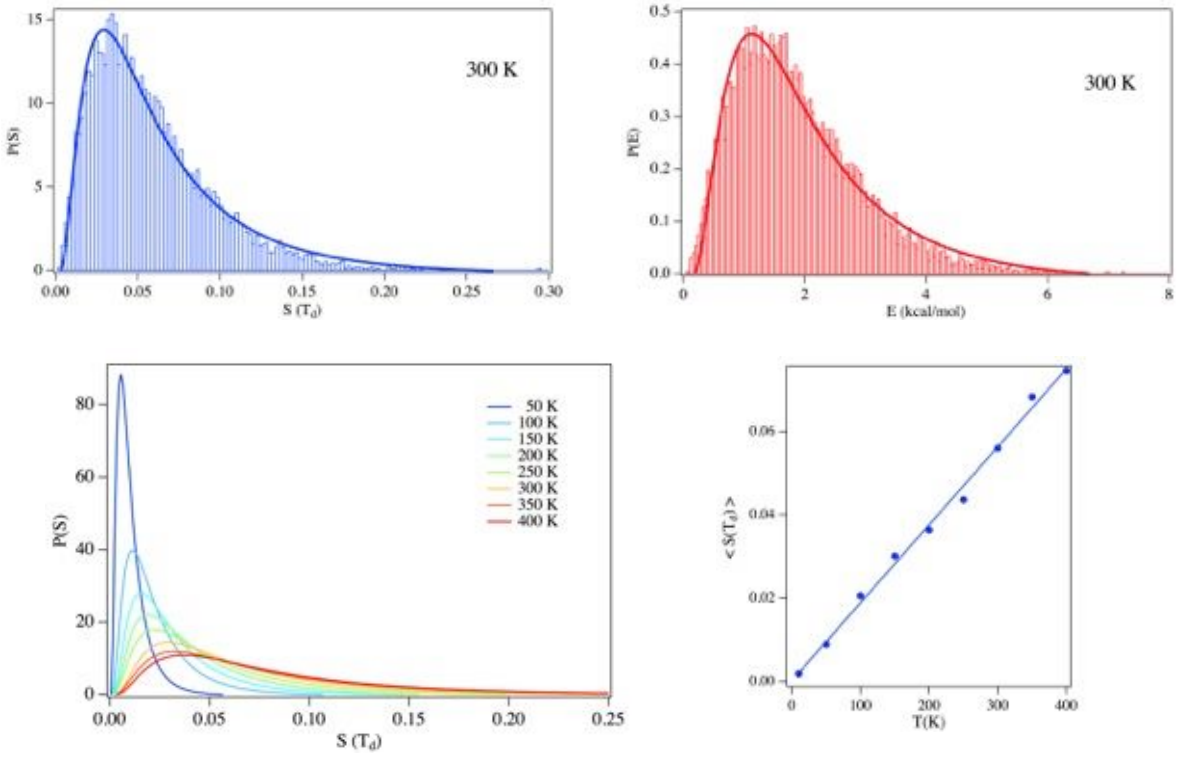


Figure 4. a) Distribution of the tetrahedral CSM for the set of 30.000 P_4 structures obtained in a simulation at T =300 K. b) Distribution of the total energy for the same set. c) Evolution for the distribution of the tetrahedral CSM with the temperature. d) Average tetrahedral CSM as a function of the temperature.

The evolution of the distribution of the CSM values with the temperature is shown in figure 4c, where it is evident that both the position of the maximum as well as the spread of the distribution increase with the temperature. An analogous behavior is found for the energy. Considering that the average value for a variable following a normal distribution is

$$\langle x \rangle = \exp\left(\mu + \frac{\sigma^2}{2}\right) \qquad [4]$$

we can easily evaluate the changes in average CMS and energy to find that both the average CSM (Figure 4d) and the average energy (not shown in the figure) increase linearly with temperature in this temperature range. From this we can conclude that there is also a linear dependence of the average CSM on the average energy. In order to better understand this result it is illustrative to plot the CSM versus the energy values for all structures in a simulation (Figure 5a).

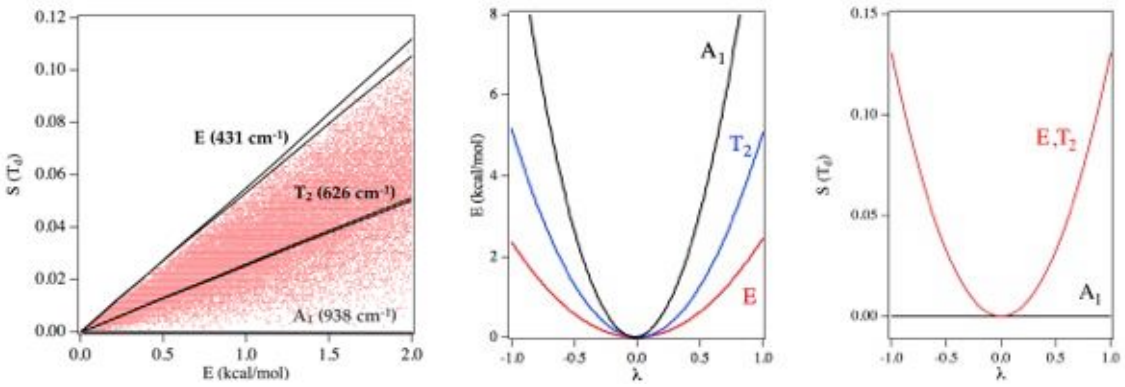


Figure 5. a) Tetrahedral CSM versus energy (red dots) for the set of 30.000 P_4 structures obtained in a simulation at T = 400K. The black lines show the variation of the CSM with the energy along the normal modes of vibration. b) Variation of the energy along the different normal modes for a single P_4 molecule. d) Variation of $S(T_d)$ along the different normal modes for a single P_4 molecule.

In this figure we see that the dots are distributed in a triangular region with one vertex at the origin, indicating that low energy structures (the minimal energy corresponds to 0) show only small deviations from tetrahedral symmetry $S(T_d) \approx 0$. Although it is not highly probable, it is possible to find perfectly tetrahedral molecules, $S(T_d) = 0$, with a high energy. To obtain these, we just need to distort the molecule along the totally symmetric A_1-type mode. As shown in figures 5b and 5c, a distortion along the A_1-type

mode implies an energy increase, but no change in the molecular symmetry contents. The more interesting feature is the upper limiting border found for the distribution of points in figure 5a. In this region of the plot is not possible (it is not the case that there are just few points with a low probability as in the lower part of the graph) to find points above a given threshold, meaning that given a certain amount of energy it is only possible to destroy the original tetrahedral symmetry up to a maximum degree. A look to figure 5a reveals that this upper border coincides, at least at low energies, with the distortion along the softest E-type normal mode. An a priori unexpected result is shown in figure 5c: except for the totally symmetric mode, the symmetry loss is exactly the same along any normal mode. It is possible to show mathematically that this result is valid for any structure with equivalent vertices,[53] leading to an interesting corollary: for a given energy, the maximum symmetry loss will be given by a deformation along the softest normal mode. As shown in figure 5a, distortions along the intermediate energy mode T_2 lead to lower symmetry losses for a same energy. A note of caution should, however, be given at this point. This general result is based on the harmonic approximation for the energy, which is only valid for small displacements around the energy minimum. In reality the variation of the energy along a given mode is not strictly parabolic, as it is assumed in the harmonic approximation, and the energetic cost is slightly different when moving along the mode in the positive or in the negative directions. This is, however, not true for the change in symmetry since in this case it can be shown that the symmetry loss is perfectly parabolic. As a consequence of this, the linear relation between symmetry loss and energy is not exactly the same when moving back or forward along a given normal mode. This effect is clearly visible in figure 5a for the E mode, the most anharmonic of the three modes, for which two slightly different lines can be observed for the change in $S(T_d)$ as a function of the energy. For small displacements, where the anharmonic effects may be neglected, the two lines coincide, marking the upper limit for the symmetry loss at a given temperature. Anharmonicity effects are smaller for the T_2 mode and the two lines are practically coincident for the displacement range shown in the figure. Evidently, for the totally symmetric mode there is no symmetry loss in any of the two directions, independently of its anharmonicity.

P_4 is an example of a molecule with a well-defined shape, a tetrahedron, robust against atomic displacements associated with temperature. The minimum on the potential energy surface corresponding to a tetrahedral geometry is well separated by high-energy barriers from those corresponding to alternative shapes such as, for instance the square. A different situation is found for flexible molecules, where various alternative shapes become accessible when the temperature increases moderately. In these cases we find low energy modes, usually corresponding to internal rotations around a single bond, which allow the switching between different conformations, which may have totally different shapes. The simplest example for this case is given by the ethane molecule, where internal rotation around the central C-C bond allows the exchange between different staggered (D_{3d}) conformations with a low barrier, aprox. 1 kcal mol^{-1} at the PM6 level, corresponding to an eclipsed (D_{3h}) conformation (Figure 6a). As far as the shape is concerned, the most stable staggered configuration may be seen to correspond to an octahedron elongated in the direction of one of its C_3 symmetry axes that would coincide with the direction of the C-C bond. Changing the H-C-C-H dihedral angle ϕ from the 60º value corresponding to the staggered geometry to 0º or 120º corresponding to the eclipsed one, can be interpreted as a Bailar-

type distortion,[23c] where one triangular face of the octahedron is rotated around the C_3 symmetry axis perpendicular to it while the opposite face remains in its original position. The result after a 60º rotation is a triangular prismatic geometry, which is precisely the geometry found for the six hydrogen atoms in the eclipsed conformation of ethane.

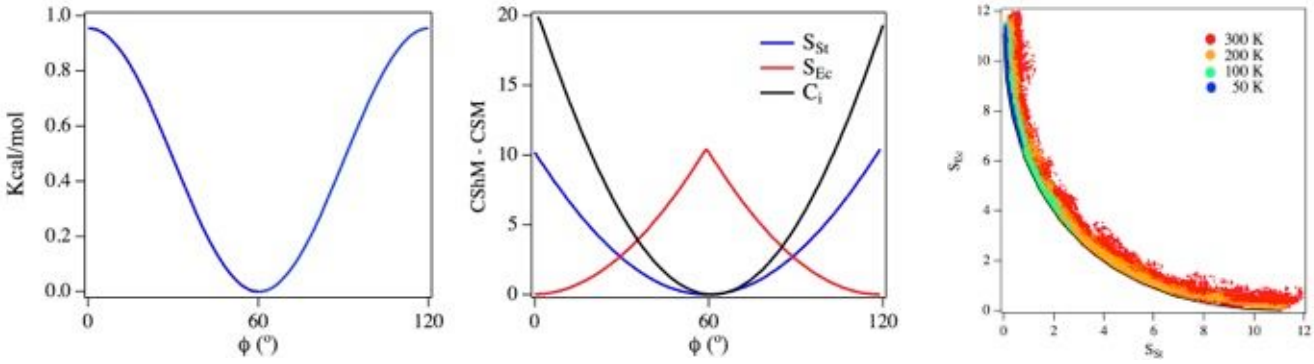


Figure 6. a) Potential energy curve for the internal rotation in ethane. b) S_{St}, S_{Ec} and S_{Ci} as a function of the dihedral angle c) Distribution of CShMs for ethane plotted in an alternate/staggered shape map. For each temperature a sequence of 60.000 structures have been generated using a Monte Carlo algorithm in which the Cartesian coordinates for all 8 atoms in the molecule are allowed to change, with the energy being evaluated at the PM6 level.

The question on how the temperature affects molecular shape is here trickier, since we have two alternative choices to describe it. We could choose to take the octahedron and the trigonal prism as reference polyhedra, but a more accurate description for this case is to take as arbitrary reference shapes those obtained for the molecule optimizing the geometry at a fixed ϕ = 60º or 0º value, that we will designate as staggered (St) and eclipsed (Ec) shape, respectively. An alternative way of characterizing the molecular shape in this case is to use the inversion symmetry measure. Since C_i is a subgroup of D_{3d}, the symmetry group for the staggered conformation, but not of D_{3h}, the group corresponding to the eclipsed one, we can use the inversion measure as a single parameter to characterize the changes in shape/symmetry of the ethane molecule as a function of the temperature. For this purpose it is useful to look at how the C_i CSM changes along the minimum energy path joining the staggered and the eclipsed conformations (figure 6b).

As a first approximation we can consider that only the lowest energy mode, corresponding to the internal rotation, is active and the only accessible distortions are coincide with those molecular geometries corresponding to the relaxed energy versus dihedral angle curve shown in figure 6a. For very low temperatures we expect the internal rotation to behave as an oscillation around the minimum energy configuration with ϕ = 60º and the best description for the molecular shape will be given by the staggered shape with $\langle S_{St} \rangle \approx 0$, $\langle S_{Ec} \rangle \approx 10$, and $\langle S_{Ci} \rangle \approx 0$, the values corresponding to the staggered geometry (Figure 6b). If we plot the S_{St} and S_{Ec} measures in a shape map (Figure 6c) we find that at low temperatures the whole population is concentrated in the surroundings of the position of the staggered shape at the point (0,10). When the temperature is raised, the population spreads along the minimum

distortion path between the two ideal shapes, reaching eventually the region corresponding to the eclipsed structure (point (10,0) on the shape map) if the temperature is sufficient to allow the system to jump over the barrier. Since in our simulations we allow al 8 atoms in the molecule to move freely, the structures generated along the Monte Carlo simulation are not constrained to lie exactly along the minimum distortion path, although the spread away from this path is not very large, indicating that the model considering only distortions along the minimum energy path is quite appropriate at the temperatures considered in the simulations.

Considering the behavior described previously for P_4 we expect $\langle S_{St} \rangle$ and $\langle S_{Ci} \rangle$ to increase linearly as the temperature rises and this is indeed what is found in the simulations for very low temperatures, just below 50K. The average CShM for the eclipsed geometry, $\langle S_{Ec} \rangle$ is expected to decrease consequently. The linear behavior is, however lost above 50K, where $\langle S_{St} \rangle$ and $\langle S_{Ci} \rangle$ have a more pronounced increase with the temperature, an increase that is being progressively moderated at even higher temperatures. Extrapolating the curves at very high temperatures, around 4000K, where any value for the dihedral angle would have practically the same probability, we would finally reach $\langle S_{St} \rangle \approx 3.7$, $\langle S_{Ec} \rangle \approx 3.4$, and $\langle S_{Ci} \rangle \approx 6.6$, the average values corresponding to a uniform distribution of dihedral angles. This situation is, however, never reached since besides the fact that the molecule would probably dissociate at such high temperatures, at intermediate temperatures the assumption that the only source of distortion is due to the internal rotation mode does not hold anymore and we would observe additional shape and symmetry losses in the simulations due to the presence of molecular geometries that are no longer restricted to those on the potential energy curve shown in figure 6a. For C_2H_6 the normal mode corresponding to the internal rotation calculated at the PM6 level appears at $\nu = 180$ cm^{-1}, while the next lowest mode is a degenerate mode with $\nu = 883$ cm^{-1}. The characteristic vibrational temperatures, $\Theta_{vib} = \nu/k_B$, for these two modes are 260 and 1270 K, respectively, indicating that in the temperature range included in the figure the only mode with a significant contribution to geometrical distortion in ethane will be that corresponding to rotation around the single C-C bond.

Once we have introduced the basic notions to understand how to use the language of continuous shape and symmetry measures to analyze the effects of temperature on molecular structure we will give a brief overview of some illustrative examples that show that in many cases we might find counterintuitive results which at a first sight might seem to be in contradiction with some well accepted notions on molecular structure.

5 Polyhedral molecules

Let us consider as a first example the set of highly symmetric C_4H_4, $B_6H_6^{2-}$, C_8H_8, $B_{12}H_{12}^{2-}$, and $C_{20}H_{20}$ molecules (Figures 7a-e) with a fairly rigid core formed by main-group atoms adopting the shape of the five platonic solids, the tetrahedron, octahedron, cube, icosahedron, and dodecahedron, respectively, in their minimum energy configurations. In this case we will center our attention first at the symmetry loss experienced by the central core of these molecules when the temperature is raised. Pertinent questions

to answer are: a) will the central core of the molecule loose its shape (symmetry) at the same rate as a simple cluster with the same shape such as the tetrahedral P_4 molecule discussed above, or is there any noticeable effect of the surrounding H atoms? b) is the number of symmetry operations in the symmetry group (24 for T_d, 48 for O_h, and 120 for I_h) a relevant parameter to describe the average symmetry loss? or in other words, is it easier to loose symmetry for a structure belonging to a high symmetry group than to a lower symmetry one? and, in general, c) which are the molecular features that are important to describe the rate of symmetry loss for this family of compounds?

Simulations for these five polyhedral molecules show again that the distributions of energy and polyhedral shape for the central cluster are again following a log-normal distribution. As shown in figures 7f and 7g, the increase in average energy or in average polyhedral shape with temperature follow both a nice linear behavior.[54] It is however evident that while the effect of temperature on the average energy follows a simple trend, that is, the larger the molecule, the more increases the average energy as the temperature is raised, this is not longer true for the effect of the temperature on the average shape loss. In this case, the least affected molecule is dodecahedrane, the largest one, although the trend is not trivial and octahedral $B_6H_6^{2-}$ is the structure most affected by temperature

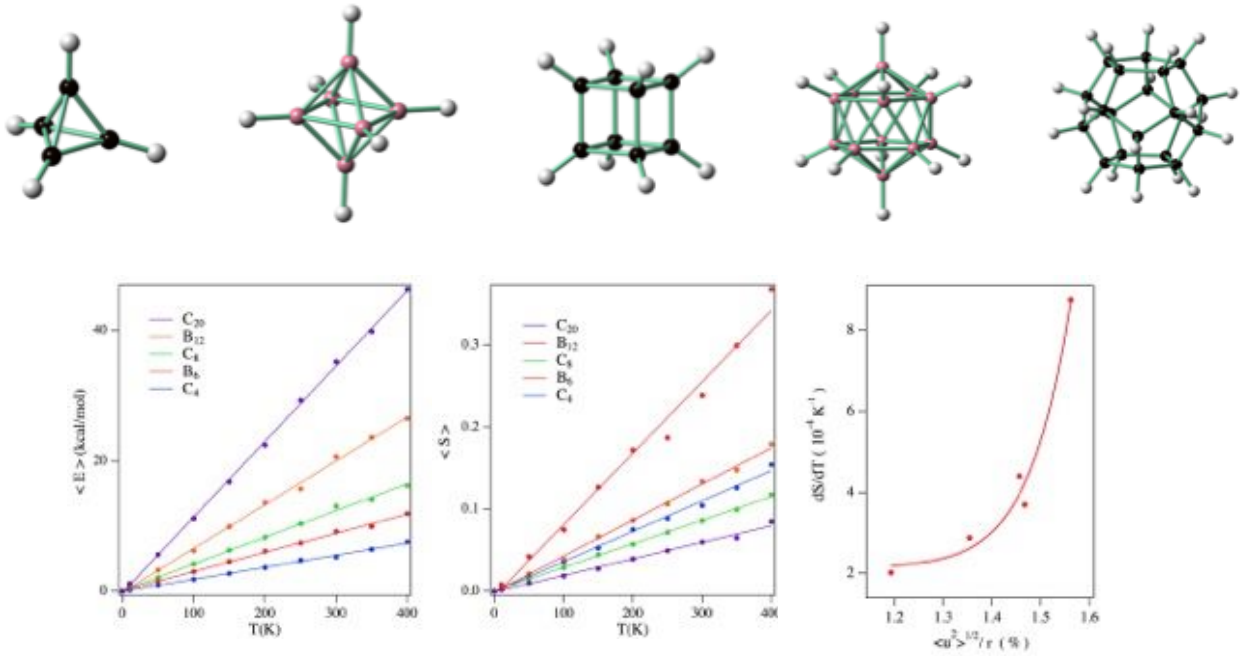


Figure 7. Molecular structures for (a) C_4H_4, (b) $B_6H_6^{2-}$, (c) C_8H_8, (d) $B_{12}H_{12}^{2-}$, and (e) $C_{20}H_{20}$ with tetrahedral, octahedral, cubic, icosahedral, and dodecahedral main group atom cages. Average energy (f) and average CSM (g) as a function of temperature for the series of C_4H_4, $B_6H_6^{2-}$, C_8H_8, $B_{12}H_{12}^{2-}$, and $C_{20}H_{20}$ molecules. (h) Slope of the average CSM vs T curves as a function of the relative atomic displacement at room temperature.

The behavior observed for the variation of the average energy with temperature (Figure 7f) can be easily rationalized if we take the equipartition theorem into account. According to this principle we should expect:

$$\langle E \rangle = \frac{3}{2} N k_B T \qquad [5]$$

where N is the total number of atoms (including the H ones) in the molecule, a trend that is in great measure the behavior found in our simulations. As it was expected, the average energy for each atom to move away from its equilibrium position is thus similar for all five molecules and the trends found for the symmetry loss must be due to some other reasons.

A simple explanation for the behavior shown in figure 7g can be found by considering eq. [2]. As it was mentioned above, the loss of shape is not directly related to the average displacement of the atoms from their equilibrium positions, but to the magnitude of this average displacement with respect to the total size of the structure. The main variables involved in eq. [2] are the masses of the atoms and the rigidity of the structure, described by the frequency of the vibration. In our case the difference in masses between B and C is not so much relevant, but the rigidity of the different molecules is quite different. If we consider the frequency of the totally symmetric breathing mode for the inner cluster as an average measure for the rigidity of the structure ($\omega \approx$ 51, 39, 38, 30, and 27 THz for C_4H_4, $B_6H_6^{2-}$, C_8H_8, $B_{12}H_{12}^{2-}$, and $C_{20}H_{20}$, respectively) we will find that the average displacement will be lowest for the most rigid tetrahedrane molecule and largest for dodecahedrane. It is also noticeable that, in general, hydrocarbons are more rigid than the boranes, for instance the smaller $B_6H_6^{2-}$ has practically the same frequency for the breathing mode than C_8H_8. If we consider also the radii of the clusters (r = 0.92, 1.20, 1.36, 1.66. and 2.17Å for C_4H_4, $B_6H_6^{2-}$, C_8H_8, $B_{12}H_{12}^{2-}$, and $C_{20}H_{20}$, respectively) as a measure of their size, the ratio between the size of the average displacement calculated using eq. [2] and the radius of the cluster gives us the clue to understand the trends found for the symmetry loss as the temperature is raised. In figure 7h it can be appreciated that the larger the ratio between the average displacement and the radius, the larger the slope of the <S> vs. T curve, showing that it is the combination of rigidity and size of the cluster that are making the $B_6H_6^{2-}$ molecule the most susceptible of all five against temperature induced symmetry loss. This reasoning explains also why boranes are more prone to loose symmetry with temperature than hydrocarbons since they are, in comparison with similar sized hydrocarbons, less rigid. For the three hydrocarbons we find that temperature is affecting much more the symmetry in tetrahedrane than in cubane and dodecahedrane. Note that the trends in the magnitude of the average symmetry loss are not related to the number of operations in the symmetry group, so that the intuitive idea that it should be easier to loose symmetry in a structure with a higher symmetry does, in principle, not hold.

By comparing the behavior of C_4H_4 with that of P_4 described previously we find that the presence of the outer hydrogen atoms has no significant effect on the evolution of the shape or symmetry of the inner core with temperature and that the shape of the carbon core in tetrahedrane is much more affected by temperature than in the P_4 molecule, an effect that can be adequately rationalized using again eq. [2] and

taking into account that the C_4 core in tetrahedrane is smaller and the C atoms lighter than the P ones, although in compensation the C_4 is somewhat more rigid than the P_4 one.

A last interesting question for these polyhedral A_nH_n molecules with an external H_n polyhedron that has the same shape as the inner one, is to explore if temperature is affecting the shape of the two polyhedra in the same amount or not. Since the H_n polyhedron has a larger radius, we expect that for a similar average displacement of the atoms, the effect should be smaller. On the other hand, the H atoms are lighter than either the B or C atoms and the totally symmetric C-H or B-H stretching mode has a much higher frequency, so that to a good approximation we could consider that H atoms will "follow" the movement of the B or C atoms in the inner framework. For this reasons it is difficult to make a qualitative prediction of which factor will dominate and which symmetry loss will be larger or smaller, that of the inner polyhedron or that of the outer H one. The answer is that according to the CSM's, although the symmetry loss increases linearly for both polyhedra as the temperature is raised, the effect on the outer H shell is larger than for the core main-atom polyhedron. For tetrahedrane, for instance, at 400K the average symmetry loss for the C_4 tetrahedron is about 0.15 while that for the outer H_4 tetrahedron is 0.29, almost two times larger. An interesting question is if the symmetry loss in the two polyhedra is related or not, a possibility that is suggested by the fact that due to the high frequency C-H vibrations one might consider that the position of the H atom could be in some way dependent on the movement of the underlying C atom. The answer to this question is, however, that for the samples obtained in the simulation, the two symmetry losses do not show any correlation and, at least for the range of temperatures explored, the atomic motions in the inner and the outer shells seem to be practically independent.

6 Intrinsic versus extrinsic symmetry loss

Up to now, we have discussed the effect of temperature on for molecules with a highly symmetric equilibrium configuration. An interesting question, raised by Tuvi-Arad and coworkers,[52] is if it is possible to distinguish the extrinsic symmetry loss due to atomic motion that appears on raising the temperature from an intrinsic lack of symmetry present already in the equilibrium geometry. A nice example to illustrate this problem can be found for some unsaturated cyclic hydrocarbons, which are known to adopt regular polygonal structures when their electronic structure satisfies Hückel's 4n+2 aromaticity rule. Benzene, for instance, in its ground state has a perfect D_{6h} symmetry with the inner C_6 ring having the shape of a regular hexagon. If we consider the effects of temperature on benzene's shape, a picture very similar to that described above for polyhedral molecules emerges. At a given temperature, the CSM for the regular hexagon shows a log-normal distribution with the average D_{6h} CMS increasing linearly with temperature.

A more interesting case is that of cyclopentadiene, C_5H_5. The neutral molecule is one π electron short to fulfill Hückel's rule and, accordingly, it is susceptible to a Jahn-Teller distortion that lowers its ground-state symmetry from D_{5h} to C_{2v} (Figure 8a). Adding an extra electron as in the cyclopentadienyl anion restores the pentagonal symmetry for the ground state. It is quite instructive to compare the distributions for the D_{5h} (pentagonal) symmetry measures for these two structures and look at how these evolve as the temperature is raised (Figures 8b and 8c).

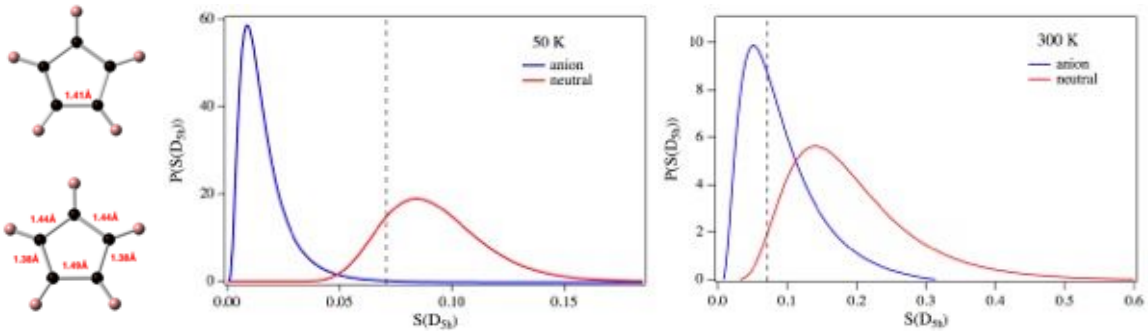


Figure 8. a) Minimum energy structures for the D_{5h} cyclopentadienyl anion (above) and neutral C_{2v} cyclopentadiene (below). b) D_{5h} CSM distributions at 50K for neutral and anionic cyclopentadiene. c) D_{5h} CSM distributions at 300K for neutral and anionic cyclopentadiene. The dashed line in figures 8b and 8c indicates the $S(D_{5h})$ value for the C_{2v} minimum energy structure found for the neutral cyclopentadienyl molecule.

For the cyclopentadienyl anion, with a perfectly D_{5h} symmetric ground-state structure, we find a behavior totally analogous to that described in the section above. The distribution for the pentagonal shape measure has a log-normal behavior (blue curves in figures 8b and 8c), with a sharp increase at low CSM values, and after reaching a maximum, a more gentle descent towards zero. This peculiar asymmetric shape of the log-normal distribution is a consequence of the fact that CSMs are always positive. Any distortion from the initially pentagonal structure will lead to a positive value for $S(D_{5h})$. At low temperatures we have the sharp peak close to zero, while increasing the temperature results in a displacement of the peak towards higher CSMs values and a broadening of the peak that leads to a less asymmetric curve. The effects of the temperature on the distribution may be described by the changes in the two parameters, μ and σ, describing the log-normal probability distribution function (eq. [3]), or alternatively by giving the mode (position of the maximum) and the skewness of the distribution, which for the log-normal distribution are related to μ and σ as:

$$Mod(P) = e^{\mu-\sigma^2}$$
$$Skw(P) = \left(e^{\sigma^2}+2\right)\sqrt{e^{\sigma^2}-1} \qquad [6]$$

For the cyclopentadienyl anion, with a regular pentagonal shape for the inner C_5 ring, the mode of the distribution shifts from 0.01 at 50 K to 0.05 at 300K, while the skewness is reduced from 1.17 at 50K to 0.55 at room temperature.

Let's compare this behavior with that for the non-aromatic neutral C_5H_5 molecule. Geometrical optimization with the PM6 method leads to a C_{2v} symmetry with two short, two intermediate and a long C-C bond (figure 8a). Departure from the pentagonal symmetry, $S(D_{5h})$ = 0.071, is not large, but chemically significant and of the same order as temperature induced symmetry losses for the symmetrical cyclopentadienyl anion. If we look at the evolution of the distribution of shape measures as a function of

the temperature (red curves in Figures 7b and 7c) we find some striking differences with those found for the cyclopentadienyl anion. Since now the lowest energy structure is not the most symmetrical one, as the temperature rises, the displacement of atoms may lead to less symmetric structures, but also to more symmetric ones, so that at a finite temperature the distribution function for the shape measure may extend both to the right and to the left of the CMS value corresponding to the equilibrium structure (indicated by a dashed line in Figures 8b and 8c). The effect of the possibility of symmetrizing distortions translates into a much less skewed distribution, as it is clearly evident at 50 K (Figure 8b). Note, however, that the mode of the distribution, that is the most probable CMS value is 0.083, slightly larger than the CMS corresponding to the equilibrium structure, 0.071. The exact position of the maximum of the distribution with respect to the symmetry measure is difficult to predict since it depends on the details of the potential energy surface and the relative strength of the normal mode leading from the pentagonal to the C_{2v} structure. The effect of increasing the temperature is again to shift the peak of the distribution to higher values, from 0.083 to 0.142 in this case when we go from 50 to 300K, increasing the average symmetry loss as in the case of symmetrical molecules. What is interesting to note is, that since temperature shifts the average symmetry measure to higher values and at the same times results in a broadening of the distribution, in this case the effect on the asymmetry of the log-normal distribution is exactly the opposite to that found for the symmetric molecule: the distribution's skewness now increases, from 0.55 at 50K to 1.02 at 300K. The changes in the average symmetry measure show also a linear increase with the temperature. The slope for this variation is very similar to that found for the symmetric anion, although slightly larger. If one plots the evolution of the average CSM for both cases in the same plot, one finds two practically parallel lines, the one for the symmetric cyclopentadienyl anion with a CSM value of 0 at 0K (if the symmetry loss due to zero point vibrational movement is neglected) and that for the neutral compound starting at a 0.071 value at 0K, the pentagonal CSM corresponding to the equilibrium C_{2v} structure.

A similar situation can be found for cyclobutadiene: for the $C_4H_4^{2-}$ dianion satisfying Hückel's rule we find a behavior totally analogous to that described for the cyclopentadienyl anion or for benzene, with the minimum energy geometry corresponding to a perfect C_4 square. For the neutral compound with two electrons in a doubly degenerate molecular orbital set we have now two different options. If we consider a triplet ground state as predicted by Hund's rule, we find a also square minimum energy geometry, leading to distributions of square CSM values very similar to those found for the dianion. If we consider, however a singlet ground state, energy minimization leads to a Jahn-Teller distorted rectangular structure with an alternation of short and long C-C bonds. The distributions of square CSM values behave in this case in the same way as those for the neutral cyclopentadiene molecule, with the skewness of the distribution increasing as temperature is raised.

7 Effects of temperature on structural chirality

The two examples discussed above correspond to molecules with fairly rigid polyhedral or polygonal frameworks. We have seen in the introduction that in flexible molecules, with vibrational modes corresponding to internal rotations such as in ethane, the situation is more complex and the effects of temperature on symmetry or shape depend largely on the details of the potential energy surface. Let us

consider as an example for this case by examining how temperature affects the chirality of the biphenyl molecule ($C_{12}H_{10}$). The PM6 minimum energy geometry for the biphenyl molecule is found for a φ = 58º dihedral angle between the two phenyl rings (Figure 9a). This minimum energy configuration is separated from its mirror image at φ = 122º by a low energy barrier (0.2 kcal/mol at the PM6 level) at φ = 90º when the two phenyl rings are perpendicular to each other. A larger barrier (1.8 kcal/mol) at φ = 0º and 180º, when the two phenyl rings are on the same plane, separates the low energy region around φ = 90º with its equivalent centered around φ = 270º. As shown in figure 9b, except for the configurations with perpendicular (φ = 90, 270º) or coplanar phenyl (φ = 0, 180º) rings all other structures are chiral (CCM $\neq$ 0), with enantiomeric pairs, called in this case atropoisomers, at each side of the barriers. If we consider a rigid torsion around the central C-C bond, that is, if the geometry of the rings is kept fixed when changing the dihedral angle, the maximum chirality is obtained for a geometry in the midpoint between the two achiral structures, that is at φ = 45º. In the case that the geometry of the phenyl rings is allowed to relax for each given dihedral angle as in figure 9b, the maximum for the chirality is slightly displaced, around φ = 43.5º in the curve obtained using the PM6 method. This small difference does however not invalidate the interpretation of the results below if we consider that the chirality maximum is exactly at φ = 45º.

Two aspects are noteworthy in figure 9b. The first one is that the qualitative behavior for the CCM is exactly the same if we consider just the atoms in the carbon skeleton or if we take into account also the exterior hydrogen atoms. Since due to its combinatorial nature the calculation of the CCM might become heavy when the number of atoms in a structure increases, it is usual to restrict the calculation of CSMs to the main atom framework, neglecting the hydrogen atoms. As seen in the example of biphenyl, the qualitative results are the same and if we want to compare, for instance, x-ray structures for different molecules we will avoid with this procedure uncertainties arising from the difficulty in determining the position of the hydrogen atoms that could introduce spurious changes in the CSMs. The other interesting point to note is the cusp appearing in the CCM curve for the most chiral structure. This behavior, which is often found when plotting CSMs as a function of a structural parameter, originates naturally from the very definition of the CSM, eq. [1]. If we recall this definition, the chirality (or in general symmetry) measure is defined as the distance between our problem structure and the closest achiral structure. In our case there are two alternative chiral structures, the coplanar structure at φ = 0º and the orthogonal one at φ = 90º. For any structure with 0º < φ < 45º the closest achiral structure taken as a reference to calculate the CSM is the coplanar structure, while for structures with 45º < φ < 90º it will be the orthogonal one. The maximum chirality structure is equidistant from the two reference achiral structures and it is the switch between reference structures in the calculation of the CSM at this point that gives rise to the singularity in the CCM curve.

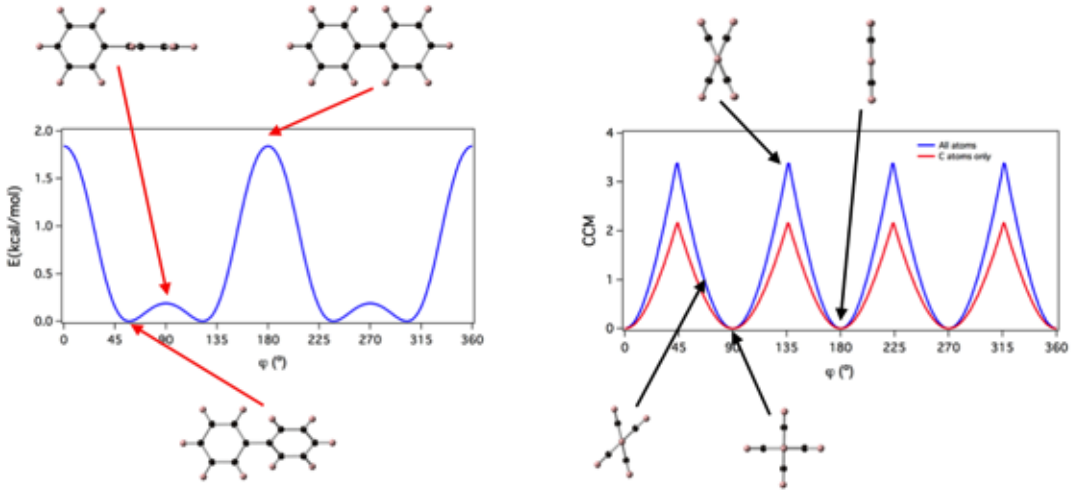


Figure 9. a) Potential energy curve for the biphenyl molecule along the internal rotation path. b) Variation of the continuous chirality measure (CCM) for the biphenyl molecule along the internal rotation path. The CCM has been calculated considering the coordinates of all atoms in the molecule (blue curve) and only those corresponding to the carbon atoms (red curve).

The joint analysis of the energy and symmetry curves allows us to predict that the evolution of chirality with temperature will not be as simple as in the previous examples. At low temperatures, where the low energy barrier at φ = 90º can still not be crossed, we will find the population of molecules with dihedral angles oscillating around the minimum energy configuration at φ = 58º with a corresponding CCM of about 1.1 if we consider only the carbon atom framework. Since the potential energy curve is somewhat asymmetric around φ = 58º, with higher φ values corresponding to lower energies, the average CCM will initially decrease with the temperature. When the temperature is sufficient to pass the barrier at φ = 90º, the dihedral angle will oscillate around this value and the molecules will be able to switch between the two atropoisomers. This does, however, not lead to a significant change in the symmetry measure since the two atropoisomers have exactly the same chirality measure. Increasing the temperature spreads the population of molecules with dihedral angles around 90º. Since the CCM starts to decrease for geometries below φ = 45º or above φ = 135º, in general, the average CCM will descend with temperature until a temperature is reached in which the barrier at φ = 0º can be overcome. From this point onwards the distribution will become more and more uniform (each dihedral angle will have similar probability) and the average CCM will tend to the average CCM value along the curve, CCM = 0.78 if we consider only the C framework. Let us recall, however, that this reasoning is based just using the Maxwell distribution for molecules along the relaxed energy versus φ curve. At low temperatures this is a good approximation, but at higher temperatures the participation of other normal vibrational modes to the atomic displacement will lead to distorted geometries with higher CCM values, and in general to an increase in the average chirality with temperature. This increase is, however, predicted to be linear as found for the examples discussed previously. The exact behavior for the CCM at low temperatures, calculated considering only geometries along the relaxed energy versus φ curve, is shown in figures 10a for two different temperatures and the evolution of the average CCM with the temperature up to 100K is shown in figure

10b. As it is evident from these figures, the distribution of CCM values at a given temperature once the barrier at φ = 90º can be passed is far from being described by a log-normal curve and the evolution of the average CCM with temperature is far from being described by a simple linear relation as in the case of a potential energy curve with a single well defined energy minimum well separated in energy from other minima by large potential energy barriers. In this sense the effects of temperature on shape and symmetry for highly flexible molecules will be, in general, complex and a detailed knowledge of the particular potential energy surface and of the variation of the CSM with geometry will be needed for a reasonable explanation of these effects.

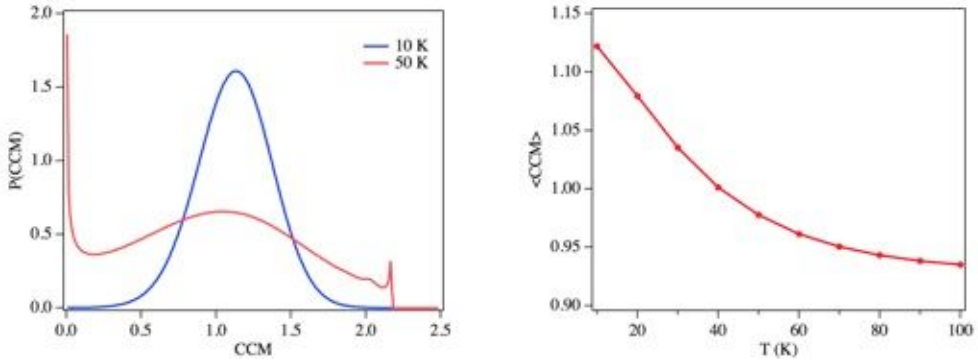


Figure 10. a) Distribution of the continuous chirality measure (only C framework) for biphenyl at 50 and 100K b) Average CCM as a function of the temperature.

8 Tetrahedral solids

In this last example we will explore how local symmetry in a crystal structure is progressively lost as the temperature is raised. The cubic diamond-type structure is the prototypical example of a crystalline structure for tetrahedral solids such as carbon, silicon, or germanium. In this polymorph all atoms in the unit cell are equivalent, with a perfect tetrahedral coordination environment in the equilibrium structure. Upon an increase of the temperature, the chaotic displacement of atoms around their equilibrium position is responsible for breaking the local tetrahedral symmetry around each single atom in the lattice. In order to analyze the effects of temperature on the shape of the coordination environments, we have run several molecular dynamics simulations[55] at different pressure and temperature values using a 512 atom super-cell for the diamond-type structure of C, Si, and Ge. As expected, the shape of the coordination environment is basically affected by temperature with the average tetrahedral shape measure increasing linearly as the temperature is raised (Figure 11a). The effect of pressure is much smaller and will not be commented here in detail, but the basic outcome is that increasing the pressure slightly inhibits the symmetry loss by increasing the structure's rigidity.

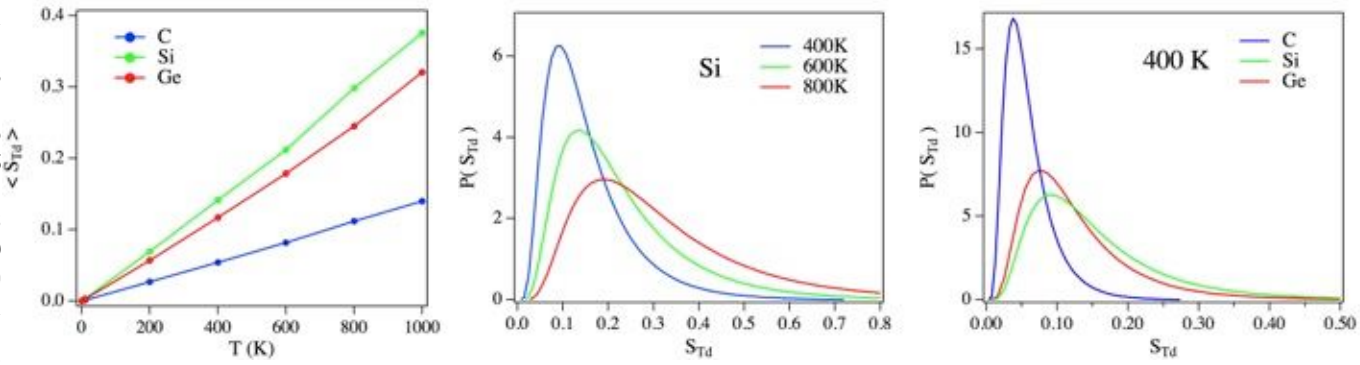


Figure 11. a) Average tetrahedral CSM for C, Si, and Ge as a function of temperature. b) Distribution of tetrahedral CSM values for Si at different temperatures. c) Distribution of tetrahedral CSM values for C, Si, and Ge at T = 400K.

At a given temperature, the probability for the tetrahedral CSM of the coordination environment around each atom is found again to follow a log-normal distribution law, with its maximum shifting towards higher values as the temperature increases (Figure 11b). If we compare the probability distributions obtained for C, Si, and Ge at a same temperature (Figure 11c), we find that, as expected, in the case of the more rigid carbon structure the distribution is much narrower, with its peak located at lower CSM values than for Si and Ge. For these two elements we find similar distributions, although unexpectedly, it seems that the symmetry loss is somewhat larger in Si than in Ge, although intuition tells us that the silicon structure should be more rigid than the germanium one. This fact is confirmed by calculating the phonon spectrum for the three structures with the same force field used in the simulations. If we take the frequency of the transverse optical mode at the center of the Brillouin zone as a qualitative measure of the crystal structure's rigidity, we find that $\nu_{TO}(\Gamma)$ = 48, 16, and 10 THz for C, Si, and Ge, respectively, indicating that carbon has, indeed, the most rigid crystal structure, for which we expect the smallest displacements of atoms and the smallest symmetry loss with temperature, as it is actually found in the simulations. Silicon is, as predicted, more rigid than germanium and intuitively the effect of temperature on local tetrahedricity should be lower in silicon than in germanium, just the contrary to what is found (Figure 11a). The explanation for this apparently puzzling result is solved again if we reason in terms of eq. [2], and consider not only the dependence of the mean atomic displacement on the frequency, but include also the mass, much larger in germanium than either in carbon or silicon, as well as the relative size of the tetrahedra. While the equilibrium C-C bond distance is found to be 1.53Å, the Si-Si one is 2.35Å and the Ge-Ge one 2.45Å. Although the lower vibrational frequency in Ge hints towards larger atomic displacements as in silicon, the larger mass for Ge and the fact that the displacement has to be contemplated with respect to a larger structure in Ge, leads at the end to the right answer, that is, that the local tetrahedral symmetry is more affected by temperature in silicon than in germanium.

Another interesting question is about the degree of chirality induced in the diamond structure by thermal motion. The displacement of the atoms will give raise at each moment to distorted tetrahedra that, in general, do not preserve any of the symmetry planes, neither the S_4 improper rotation axes present in a perfect tetrahedron. In this sense, at a given temperature, the vast majority of the instantaneous tetrahedral coordination environments will be chiral and the structure will acquire an average internal chirality that, in principle, will increase with temperature. Molecular dynamics simulations for the diamond-type polymorphs of C, Si and Ge show that this is, indeed, the case: the average continuous chirality measure is not zero and it increases linearly with temperature. The effect is smallest in the rigid carbon framework and most noticeable in silicon. If at any given temperature the instantaneous structure of solids with a diamond-type lattice is chiral, why don't we observe any effects of this chirality at a macroscopic scale, for instance in some kind of optical activity? The answer is again simple if we consider the distinction between chirality and handedness. For a chiral object we may have two enantiomers, with identical structures, one mirror image of the other. If we consider a set of structures containing equal proportions of the two enantiomers, let us label them R and S, we obtain an optically inactive racemic mixture. Using the enantiomer labeling scheme proposed by Avnir and coworkers[42] for chiral distorted tetrahedra we find that, on average, at a given temperature there is always the same proportion of R and S tetrahedra, regardless of the value of the chirality measure, or in other words, the probability of generating either a left or right-handed chiral tetrahedron by applying thermal motion to an originally achiral structure is exactly the same. A practically perfect racemic mixture of tetrahedra is found for all temperatures. A more detailed analysis shows that there is not only no enantiomeric excess considering all distorted tetrahedra as a single set, but that this is also true at all different scales of distortion. If we select all tetrahedra with chiralities between two arbitrary s_1 and s_2 values for the continuous chirality measure, we will find a practically equal amount of R and S tetrahedra, regardless of the values of s_1 and s_2 that we select. This is, of course, only true if the equilibrium structure is achiral as in the diamond structure. If the local coordination is, however, chiral, small distortions will in general preserve the handedness of the coordination environment in the equilibrium structure, while large distortions may lead to a handedness switch. In these cases the enantiomeric excess for the whole crystal structure will depend on various factors such as the initial enantiomeric excess in the equilibrium structure and the temperature that is needed to have a sufficiently large atomic displacements to switch the handedness of each individual chiral center.

Summary and Outlook

Stereochemistry, the study of the relative spatial arrangement of atoms in molecules or crystals is nowadays a well established subdiscipline of chemistry, with its roots lying in the stereoformulae established in the late 19th century. The lack of experimental tools to determine precisely the atomic structure of molecules and crystals in the early development of stereochemistry lead to an abuse of the use of highly symmetrical polyhedral shapes in describing molecular structure that has been later shown to be just an idealization of more complex geometrical features. The neglect in early stereochemical studies of important factors such as the influence of the environment on molecular structure and the inherent dynamical nature of molecular structure has resulted in the widespread use of vague concepts

such as distortion by molecular packing forces, molecular flexibility or dynamical disorder that are used to explain why usually experimental data differ from the idealized models used to describe molecular shapes. In this sense the forging of a new language, that of continuous symmetry and shape measures, describing in a quantitative fashion the departure of experimentally determined molecular shapes from those supposed in the ideal stereochemical models allows chemists to deepen in the comprehension of the fundamental links between structure and the physical properties of molecules and crystals. Taking into account in a quantitative way the departure from the ideal shapes/symmetries of molecular structure induced by interactions with the environment or by temperature gives a much richer description of molecular shape that may be profitably incorporated to modern machine learning approaches to explore the complexity of structure-property relations in fields such as medicinal chemistry or materials science. The approach described in this work provides also a useful tool to analyze complex data obtained in molecular dynamics simulations, allowing to focus on the emergence of relevant structural changes in complex systems disregarding the noise of random atomic displacements not directly involved in these changes.

Conflict of interest

The authors declare no conflict of interest.

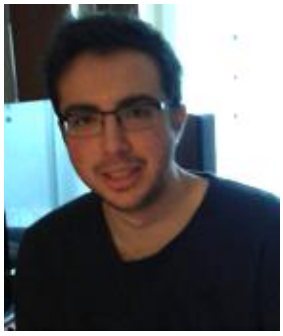
Acknowledgements

This work was financially supported by MICINN (project CTQ2015-64579-C3-3-P) and Generalitat de Catalunya (project 2017 SGR 1289). E. B. thanks the Spanish 'Ministerio de Economía y Competitividad' for a predoctoral FPI grant. The authors are specially indebted to David Avnir and Santiago Álvarez, as well as to their respective coworkers in Israel and Barcelona for the many enriching discussions on CSMs held with them during the last twenty years. P.A. would like to acknowledge the warm hospitality of Clare Hall that supported his stay at Cambridge University with a visiting fellowship during the period when this article was written.

Biographical information

Abel Carreras obtained his Ph.D. in Chemistry from the Universitat de Barcelona in 2013. He focused in the analysis of the dynamics of molecular rotors and gears. After his Ph. D. he carried out postdoctoral studies from 2014 to 2018 in the Department of Materials Science and Engineering at Kyoto University developing a computational code to calculate phonon anharmonic properties in crystals from molecular dynamics simulations.

Efrem Bernuz obtained a master's degree in Theoretical Chemistry and Computational Modelling at the University of Barcelona in 2016 working in the development of new methods to study quantum dynamics. At the present, his working towards his PhD in the group of P. Alemany where his research focuses on the analysis of the shape and symmetry in stereochemistry as well as in the bonding and reactivity of transition metal complexes.



Xavier Marugan Ferrer studied Chemistry from 2012 to 2016 at the University of Barcelona. During the degree he became interested in computational chemistry and that made him study a Master in Atomistic and Multiscale Computational Modeling in Physics, Chemistry and Biochemistry. His final works both in his undergraduate and master studies were supervised by Prof. P. Alemany were related to the study of stereochemical problems in solid state structures using the CSMs approach,

Miquel Llunell received his PhD in Chemistry at the University of Barcelona (1998). In 2000, he joined the group of Prof. R. Dovesi in Torino (Italy) to work on the computational study of electronic structure of periodic solids and the development of the Crystal program. Since 2009, he is an Associate Professor at the University of Barcelona, working mainly in the development of algorithms and computational tools for the study of molecular shape and symmetry using the continuous symmetry measures.

Pere Alemany is Full Professor in the Chemistry Department of the University of Barcelona, where he started his academic career in 1991 after a post doctoral stay at Cornell University working in the group of Roald Hoffmann. He has published about 150 research articles, mainly on the electronic structure and bonding in solids and molecules, with special focus on electronic transport and magnetic exchange coupling in low dimensional systems and the use of continuous symmetry measures to address stereochemical problems.

References

[1] P. J. Ramberg, *HYLE - International Journal for Philosophy of Chemistry* **2000**, *6*, 35-61.

[2] J. H. van't Hoff, *Voorstel tot Uitbreiding der tegenwoordig in de scheikunde gebruikte Structuur-Formules in de ruimte; benevens een daarmeê samenhangende opmerkung omtrent het verband tusschen optisch actief Vermogen en Chemische Constitutie van Organische Verbindingen*, Greven, Utrecht, **1874**.

[3] a) R. G. Woolley, *J. Am. Chem. Soc.* **1978**, *100*, 1073-1078; b) H. Primas, *Quantum Dynamics of Molecules: The New Experimental Challenge toTheorists*, Plenum Press, New York, **1980**.

[4] H. Ochiai, *HYLE - International Journal for Philosophy of Chemistry* **2013**, *19*, 139-160.

[5] S. Alvarez and J. Echeverría, *J. Phys. Org. Chem.* **2010**, *23*, 1080-1087.

[6] a) F. H. Allen, *Acta Crystallogr. B* **2002**, *58*, 380-388; b) G. Belsky, M. Hellenbrandt, V. L. Karen and P. Luksch, *Acta. Cryst. B* **2002**, *58*, 364-369; c) H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Research* **2000**, *28*, 235-242.

[7] W. H. Baur, *Acta Crystallogr. B* **1974**, *30*, 1195-1215.

[8] a) P. G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape and Topology*, VCH Publishers, New York, **1993**; b) G. A. Arteca, *Reviews in Computational Chemistry* **2007**, *1*, 191-253.

[9] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, London, **2001**.

[10] K. Mainzer, *HYLE - International Journal for Philosophy of Chemistry* **1997**, *3*, 29-49.

[11] M. Petitjean, *Entropy* **2003**, *5*, 271 - 312.

[12] a) H. Zabrodsky, S. Peleg and D. Avnir, *J. Am. Chem. Soc.* **1992**, *114*, 7843-7851; b) H. Zabrodsky, S. Peleg and D. Avnir, *J. Am. Chem. Soc.* **1993**, *115*, 8278-8289; c) H. Zabrodsky and D. Avnir, *J. Am. Chem. Soc.* **1993**, *117*, 462-473; d) M. Pinsky and D. Avnir, *Inorg. Chem.* **1998**, *37*, 5575-5582; e) H. Zabrodsky, S. Peleg and D. Avnir, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1995**, *17*, 1154-1166.

[13] a) Y. Salomon and D. Avnir, *J. Math. Chem.* **1999**, *25*, 295-308; b) Y. Salomon and D. Avnir, *J.Comp. Chem.* **1999**, *20*, 772-780.

[14] a) C. Dryzun and D. Avnir, *Phys. Chem. Chem. Phys.* **2009**, *11*, 9653-9666; b) C. Dryzun and D. Avnir, *ChemPhysChem* **2011**, *12*, 197-205.

[15] a) C. Dryzun, P. Alemany, D. Casanova and D. Avnir, *Chem. Eur. J.* **2011**, *17*, 6129 - 6141; b) P. Alemany, *Int. J. Quantum Chem.* **2013**, *113*, 1814-1820; c) P. Alemany, D. Casanova, S. Alvarez, C. Dryzun and D. Avnir, *Reviews in Computational Chemistry* **2017**, *30*, 289-352; d) E. Estrada and R. Carbó-Dorca, *Match-Communications in Mathematical and in Computer Chemistry* **2009**, *62*, 105-144.

[16] a) D. Casanova and P. Alemany, *Phys. Chem. Chem. Phys.* **2010**, *12*, 15523-15529; b) D. Casanova and P. Alemany, *Chem. Phys. Lett.* **2011**, *511*, 486-490.

[17] a) K. B. Lipkowitz, D. Gao and O. Katzenelson, *J. Am. Chem. Soc.* **1999**, *121*, 5559-5564; b) L. Bellarosa and F. Zerbetto, *J. Am. Chem. Soc.* **2003**, *125*, 1975-1979; c) P. Alemany, D. Casanova and S. Alvarez, *Phys. Chem. Chem. Phys.* **2012**, *14*, 11816-11823.

[18] D. Casanova, P. Alemany, A. Falceto, A. Carreras and S. Alvarez, *J. Comput. Chem.* **2013**, *34*, 1321-1331.

[19] A. Falceto, D. Casanova, P. Alemany and S. Alvarez, *Inorg. Chem.* **2013**, *52*, 6510-6519.

[20] D. Casanova, P. Alemany and S. Alvarez, *J. Comput. Chem.* **2010**, *31*, 2389-2404.

[21] P. Alemany, D. Casanova and C. Dryzun, *Chem. Eur. J.* **2011**, *17*, 14896-14906.

[22] a) S. Alvarez and M. Llunell, *J. Chem. Soc. Dalton Trans.* **2000**, 3288-3303; b) S. Alvarez, M. Pinsky and A. D., *Europ. J. Inorg. Chem.* **2001**, 1499-1503; c) S. Alvarez, M. Pinsky, M. Llunell and A. D., *Cryst. Eng.* **2001**, *4*, 179-200.

[23] a) S. Keinan and D. Avnir, *Dalton Trans.* **2001**, , 941-947; b) S. Keinan and D. Avnir, *Inorg. Chem.* **2001**, *40*, 318-323; c) S. Alvarez, D. Avnir, M. Llunell and M. Pinsky, *New J. Chem.* **2002**, *26*, 996.

[24] D. Casanova, P. Alemany, J. M. Bofill and S. Alvarez, *Chem. Eur. J.* **2003**, *9*, 1281.

[25] a) J. Cirera, P. Alemany and S. Alvarez, *Chem.-Eur. J.* **2004**, *10*, 190-207; b) S. Alvarez, P. Alemany, D. Casanova, J. Cirera, M. Llunell and D. Avnir, *Coord. Chem. Rev.* **2005**, *249*, 1693-1708; c) D. Casanova, M. Llunell, P. Alemany and S. Alvarez, *Chem. Eur. J.* **2005**, *11*, 1479; d) J. Cirera, E. Ruiz and S. Alvarez, *Organometallics* **2005**, *24*, 1556-1562 ; e) J. Cirera, E. Ruiz and S. Alvarez, *Chem. Eur. J.* **2006**, *12*, 3162-3167; f) A. Ruiz-Martínez, D. Casanova and S. Alvarez, *Chem. Eur. J.* **2008**, *14*, 1291-1303; g) J. Echeverría, E. Cremades, A. J. Amoroso and S. Alvarez, *Chem. Commun.* **2009**, 4242-4244; h) A. Ruiz-Martínez and S. Alvarez, *Chem. Eur. J.* **2009**, *15*, 7470-7480; i) E. Cremades, J. Echeverría and S. Alvarez, *Chem. Eur. J.* **2010**, *16*, in press; j) J. C. Knight, S. Alvarez, A. J. Amoroso, P. G. Edwards and N. Singh, *Dalton Trans.* **2010**, *39*, 3870 - 3883.

[26] a) J. Cirera and S. Alvarez, *Dalton Trans.* **2013**, 7002-7008; b) A. Falceto, D. Casanova, P. Alemany and S. Alvarez, *Chem. Eur. J.* **2014**, *20*, 14674-14689; c) S. Alvarez, *Chem. Rev.* **2015**, *115*, 13447-13483; d) S. Alvarez, *Coord. Chem. Rev.* **2017**, *350*, 3-13; e) S. Alvarez, B. Menjon, A. Falceto, D. Casanova and P. Alemany, *Inorg. Chem.* **2014**, *53*, 12151-12163; f) I. Sanchez-Lombardo, S. Alvarez, C. C. McLauchlan and D. C. Crans, *J. Inorg. Bioch.* **2015**, *147*, 153-164.

[27] a) K. M. Ok, P. S. Halasyamani, D. Casanova, M. Llunell, P. Alemany and S. Alvarez, *Chem. Mater.* **2006**, *18*, 3176-3183; b) J. Cirera, E. Ruiz and S. Alvarez, *Inorg. Chem.* **2008**, *47*, 2871; c) J. Echeverría and S. Alvarez, *Inorg. Chem.* **2008**, *47*, 10965-10970; d) J. Echeverría, D. Casanova, M. Llunell, P. Alemany and S. Alvarez, *Chem. Commun.* **2008**, 2717-2725; e) P. Alemany, I. d. P. R. Moreira, R. Castillo and J. Llanos, *J. Alloys Compd.* **2012**, *513*, 630-640; f) J. Vela, J. Cirera, J. Smith, R. Lachicotte, C. Flaschenriem, S. Alvarez and P. Holland, *Inorg. Chem.* **2007**, *46*, 60-71.

[28] a) V. Buch, E. Gershgoren, H. Helor and D. Avnir, *Chem. Phys. Lett.* **1995**, *247*, 149-153; b) S. Alvarez, *Inorg. Chim. Acta* **2010**, *363*, 4392 - 4398; c) H. Elgavi, C. Krekeler, R. Berger and D. Avnir, *J. Phys. Chem. C* **2012**, *116*, 330-335; d) H. E. Sinai and D. Avnir, *Isr. J. Chem.* **2016**, *56*, 1076-1081.

[29] a) O. Katzenelson, H. HelOr and D. Avnir, *Chem. Eur. J.* **1996**, *2*, 174-181; b) I. L. Garzón, J. A. Reyes-Nava, J. I. Rodríguez-Hernández, I. Sigal, M. R. Beltrán and K. Michaelian, *Phys. Rev. B* **2002**, *66*, 073403; c) S. Alvarez, *Dalton Trans.* **2006**, 2045-2051.

[30] a) S. Keinan and D. Avnir, *J. Am. Chem. Soc.* **2001**, *122*, 4378-4384; b) M. Bonjack-Shterengartz and D. Avnir, *Proteins* **2015**, *83*, 722-734; c) Y. Baruch-Shpigler, H. Wang, I. Tuvi-Arad and D. Avnir, *Biochemistry* **2017**, *56*, 5635-5643; d) M. Bonjack and D. Avnir, *Protein Science* **2017**, *26*, 35-36; e) M. Bonjack-Shterengartz and D. Avnir, *PLOS ONE* **2017**, *12*, e0180030.

[31] a) K. B. Lipkowitz and S. Schefzick, *CHIRALITY* **2002**, *14*, 677-682; b) P. Alemany, S. Alvarez and D. Avnir, *Chem. Eur. J.* **2003**, *9*, 1952-1957; c) S. Alvarez and A. D., *Dalton Trans.* **2003**, , 562-569; d) S. Alvarez, S. Schefzick, K. Lipkowitz and D. Avnir, *Chem. Eur. J.* **2003**, *9*, 5832-5837 ; e) S. Alvarez, P. Alemany and D. Avnir, *Chem. Soc. Rev.* **2005**, *34*, 313-326.

[32] O. Katzenelson and D. Avnir, *Chem. Eur. J.* **2000**, *6*, 1346-1354.

[33] a) S. Keinan, H. HelOr and D. Avnir, *Enantiomer* **1996**, *1*, 351-357; b) O. Katzenelson, J. Edelstein and D. Avnir, *Tetrahedron-Assymetry* **2000**, *11*, 2695-2704; c) D. Yogev-Einot and D. Avnir, *Chem. Mater.* **2003**, *15*, 464-472; d) D. Yogev-Einot and D. Avnir, *Acta Cryst. B* **2004**, *60*, 163-173; e) D. Yogev-Einot and D. Avnir, *Tetrahedron-Assymetry* **2006**, *17*, 2723-2725 ; f) C. Dryzun, Y. Mastai, A. Shvalb and D. Avnir, *J. Mater. Chem.* **2009**, *19*, 2062-2069; g) C. Dryzun and D. Avnir, *Chem. Commun.* **2012**, *48*, 5874–5876.

[34] a) D. Kanis, J. Wong, T. Marks, M. Ratner, H. Zabrodsky, S. Keinan and D. Avnir, *J. Phys. Chem.* **1995**, *99*, 11061-11066 ; b) S. Keinan and D. Avnir, *J. Am. Chem. Soc.* **1998**, *120*, 6152-6159; c) S. Alvarez, *J. Am. Chem. Soc.* **2003**, *125*, 6795-6802.

[35] A. Steinberg, M. Karni and D. Avnir, *Chem. Eur. J.* **2006**, *12*, 8534-8538.

[36] a) D. Casanova, P. Alemany and S. Alvarez, *Angew. Chem. Int. Ed.* **2006**, *45*, 1457-1460; b) A. Ruiz-Martínez, D. Casanova and S. Alvarez, *Chem. Eur. J.* **2010**, *16*, 6567 - 6581; c) I. Tuvi-Arad and D. Avnir, *J. Math. Chem.* **2010**, *47*, 1274-1286; d) I. Tuvi-Arad and D. Avnir, *J.Org. Chem.* **2011**, *76*, 4973 - 4979; e) I. Tuvi-Arad and D. Avnir, *Chem. Eur. J.* **2012**, *18*, 10014-10020; f) I. Tuvi-Arad, T. Rozgonyi and A. Stirling, *J. Phys. Chem. A* **2013**, *117*, 12726-12733; g) I. Tuvi-Arad and D. Avnir, *Chem. Eur. J.* **2012**, *18*, 10014-10020; h) S. E. Canton, X. Zhang, M. L. L. Daku, Y. Liu, J. Zhang and S. Alvarez, *J. Phys. Chem. C* **119**, *119*, 3322-3330.

[37] a) Y. Pinto, H. HelOr and D. Avnir, *J. Chem. Soc. Faraday Trans.* **1996**, *92*, 2523-2527; b) Y. Pinto, P. Fowler, D. Mitchell and D. Avnir, *J. Phys. Chem. B* **1998**, *102*, 5776-5784; c) Y. Pinto, Y. Salomon and D. Avnir, *J. Math. Chem.* **1998**, *23*, 13-29; d) M. Llunell, P. Alemany and J. M. Bofill, *Theor. Chem. Acc.* **2008**, *121*, 279 - 288.

[38] a) G. Brancato and F. Zerbetto, *J. Phys. Chem. A* **2000**, *104*, 11439-11442; b) I. Ergaz, R. A. Toscano, G. Delgado, A. Steinberg and R. Glaser, *Crystal Growth & Design* **2008**, *8*, 1399-1405.

[39] J. Echeverría, A. Carreras, D. Casanova, P. Alemany and S. Alvarez, *Chem. Eur. J.* **2011**, *17*, 359 - 367.

[40] a) M. Pinsky, D. Casanova, P. Alemany, S. Alvarez, D. Avnir, C. Dryzun, Z. Kizner and A. Sterkin, *J. Comput. Chem.* **2008**, *29*, 190-197; b) M. Pinsky, C. Dryzun, D. Casanova, P. Alemany and D. Avnir, *J. Comput. Chem.* **2008**, *29*, 2712-2721; c) C. Dryzun, A. Zait and D. Avnir, *J. Comput. Chem.* **2011**, *32*, 2526 - 2538; d) M. Pinsky, A. Zait, M. Bonjack and D. Avnir, *J. Comput. Chem.* **2013**, *34*, 2-9; e) C. Dryzun, *J. Comput. Chem.* **2014**, *35*, 748-755; f) G. Alon and I. Tuvi-Arad, *J. Math. Chem.* **2018**, *56*, 193-212.

[41] D. Casanova, J. Cirera, M. Llunell, P. Alemany, D. Avnir and S. Alvarez, *J. Am. Chem. Soc.* **2004**, *126*, 1755-1763.

[42] D. Yogev-Einot, M. Pinsky and D. Avnir, *Tetrahedron-Assymetry* **2007**, *18*, 2295 - 2299.

[43] a) R. S. Cahn, C. Ingold and V. Prelog, *Angew. Chem. Int. Ed.* **1966**, *5*, 385-415; b) V. Prelog and G. Helmchen, *Angew. Chem. Int. Ed.* **1982**, *21*, 567-583.

[44] I. U. o. P. a. A. Chemistry, *Nomenclature of Inorganic Chemistry*, Butterworths, London, **1971**, p.

[45] M. Brorson, T. Dahmus and C. E. Schäffer, *Inorg. Chem.* **1983**, *22*, 1569–1573.

[46] a) U. Knof and A. v. Zelewsky, *Angew. Chem. Int. Ed.* **1999**, *38*, 302-322; b) O. Mamula, A. v. Zelewsky, T. Bark, H. Stoeckli-Evans, A. Neels and G. Bernardinelli, *Chem. Eur. J.* **2000**, *6*, 3575-3585.

[47] J. J. P. Stewart, *J. Mol. Model.* **2007**, *13*, 1173-1213.

[48] M. T. Dove, *Structure and Dynamics. An atomic view of materials.*, Oxford University Press, Oxford, **2003**.

[49] A. Buljan, P. Alemany and E. Ruiz, *J. Phys. Chem. A* **1997**, *101*, 1393-1399.

[50] E. Ruiz and P. Alemany, *J. Phys. Chem.* **1995**, *99*, 3114-3119.

[51] For each molecule discussed in the text we have generated a set of 30.500 structures at each temperature using a Monte-Carlo Metropolis simulation procedure where the energy for each structure has been calculated using the Gaussian 09 program[56] with the semiempirical PM6 method[47]. In all simulations we started from the equilibrium geometry and the cartesian coordinates of all atom were allowed to change to a certain amount that was adjusted in each case to obtain an acceptation to rejection rate of new configurations around 50%. The first 500 configurations in each Monte-Carlo run where discarded in the shape analysis. Since we are interested in highlighting the qualitative effects of temperature on the shape and symmetry of the molecules chosen as illustrative examples, we found that using the semiempirical PM6 method to calculate the energies of the generated structures yields a qualitative correct picture at a reasonable computational cost. Comparison of the results for small molecules with those obtained using more sophisticated methods for the energy evaluation such as the DFT based B3LYP method allowed us to confirm that the PM6 method is able to capture the qualitative behavior of shape and symmetry with temperature in a quite accurate manner. In the case of ethane, it has been necessary to increase the number of structures generated at each temperature to 60.000 in order to obtain a representative sampling of structures with different dihedral angles.

[52] I. Tuvi-Arad and A. Stirling, *Isr. J. Chem.* **2016**, *56*, 1067-1075.

[53] The use of mass weighted cartesian coordinates for the calculation of the normal modes of vibration leads to a simple relation between symmetry loss and energy in the case of structures with equivalent atoms such as P_4. For molecules with different types of atoms the different masses enter in the equations relating symmetry loss and energy change along each normal mode, and each case must be analyzed in detail to find which is the mode limiting the maximum symmetry loss at a given energy.

[54] For $B_6H_6^{2-}$ the noticeable discrepancies with a linear behavior for the average shape measure at higher temperatures are related to a poor sampling, as it may be shown by running larger simulations. We have kept, however, the data corresponding to 30.000 configurations in the figure to maintain the consistency with the rest of molecules. Problems related to poor sampling considering 30.000 structures

for the shape measure appear in this series only for $B_6H_6^{2-}$, the case where the variation of shape with temperature is largest.

[55] In the case of solid state structures, in order to get a sample of different geometries at each temperature we run Molecular Dynamics Simulations (MD) using Lammps,[57] a classical molecular dynamics program that can be used to simulate a wide range of systems from materials or soft matter to coarse-grained or mesoscopic systems. In all our simulations, we used a Tersoff potential[58] for the C, Si and Ge atoms. All simulations were carried out in the NPT ensemble. In all cases, we used supercells with 512 atoms. We took a time step of 1 fs, being the total length for all simulations 1.2 ns and the first 0.2 ns were used as equilibration steps. After the initial 0.2 ns we saved snapshots of the structure every 1000 steps. This lead to a number of approx. 51712 individual atomic environments for each simulation for which we calculated CSMs and CChMs. The number of actual atomic environments may be lower, because at high temperatures some of the atoms loose their initial four-fold coordination and were no longer considered in our study.

[56] M. J. Frisch et al., *Gaussian 09* (Gaussian, Inc., Wallingford CT, 2009).

[57] a) S. Plimpton, *J. Comp Phys.* **1995**, *117*, 1-19; b) Visit http://lammps.sandia.gov for more information.

[58] J. Tersoff, *Phys. Rev. B* **1989**, *39*, 5566-5568.