nature communications

Article

https://doi.org/10.1038/s41467-025-66054-7

# Determining molecular structure, coordination geometry, and molecular symmetry using a continuous symmetry operation measure software

Received: 10 June 2025

Accepted: 27 October 2025

Published online: 21 November 2025

Check for updates

Villads R. M. Nielsen & Thomas Just Sørensen

Chemists have a strong language describing and defining idealized polyhedra **P** and symmetry point groups **G**, but no efficient measure to correlate these to real molecular structures **Q**. The Continuous Symmetry operation Measure changes this by providing automated symmetry determination and a yardstick for quantifying deviations from symmetry. Symmetry and structure have been ascribed through experience, but this approach is error-prone and provides no measure that can correlate molecular structure to molecular properties. The Continuous Symmetry operation Measure tool solves this issue as it can quantify the symmetry of any structure that can be described as a list of points in space. Here, we compare the Continuous Symmetry Measure, the Continuous Shape Measure, and the Continuous Symmetry operation Measure approaches and demonstrate how the Continuous Symmetry operation Measure can be used as a tool to determine the molecular structure, the coordination geometry, and the symmetry of water, organic molecules, transition metal complexes, and lanthanide compounds. We conclude that the Continuous Symmetry operation Measure is not limited by any of the restrictions present in the other methods and allows a detailed analysis of e.g. phase changes and luminescence.

The atomic composition and structure determine the chemical and physical properties of any material. This statement would appear trivial, and in chemistry we always define the constitution and conformation of compounds. However, the actual structure—which we refer to as the molecular structure—is frequently generalized broadly using hybridization in organic chemistry and coordination polyhedra in inorganic chemistry and materials science. Often, this generalization is lacking. The electronic structure and the chemical properties are determined by symmetry, and we propose that new tools now allow chemistry to move from generalized structures to exact molecular structures of identified symmetry[1–8].

In chemistry, symmetry considerations are common. Electronic structures of organic molecules are simplified using group theoretical considerations. Similarly, the intricacies of metal complexes are reduced using symmetry in the framework of crystal field and ligand field theory. This allows electronic transitions to be calculated and experimental spectra to be assigned[9–12]. Further, symmetry dictates whether electronic transitions are allowed or forbidden[13,14], which in turn enables the design of molecules with intricate photophysical properties[15–19]. In general, symmetry reduces the complexity of structure determination, and can be used as a decisive argument when creating structure-property relationships.

To identify and determine symmetry, we must agree on several premises. First, we must lock an axis for each symmetry under consideration[20]. Second, we must determine how much of a compound that must be included in the molecular structure, for the

Nano-Science Center & Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 København Ø, Denmark. e-mail: TJS@chem.ku.dk

 https://doi.org/10.1038/s41467-025-66054-7

molecular structure to be true[21]. And finally, we must agree on a yardstick for determining symmetry and then start using it. Here, we investigate symmetry in the molecular structure of water, polycyclic aromatic hydrocarbons (PAHs), classical transition metal complexes, single molecule magnets, and lanthanide complexes. The level of detail that we go to has recently been reviewed with a focus on molecular conformations and polymorphs[22], where our focus is on the minute structural details that determine the properties of lanthanide complexes. Here, we illustrate this using neodymium(III) and europium(III) electronic structure and luminescence.

## Defining structure

We define a molecular structure as a set of atomic coordinates that gives rise to a specific electronic structure. In some cases, the electronic structure is maintained across variations in the atomic coordinates[23–25]. In other cases, minute changes in atomic coordinates changes the electronic structure, and each set of atomic coordinates is a separate molecular structure. A molecular structure thus by definition gives rise to a molecular entity as per the IUPAC Gold Book. Do note that a molecular entity can have several molecular structures[23].

The distinction between molecular entities may appear ethereal, but in fact it is a highly relevant across chemistry. Structure-property relationships require determination of distinct structures. Seemingly simple molecules such as water $\mathbf{H_2O}$, become complex when more than a few molecules are involved[26,27]. Agglomerates of water molecules form different molecular entities with distinct properties[28,29], and the complexity is increased further in bulk water $\mathbf{H_2O}$[30–34], and if water interacts with surfaces[35,36]. Even in the solid state, the individual water molecules form several molecular entities that are different across the various forms of ice[37–39]. Each molecular entity is defined by symmetry.

Similarly, the structure of nanotubes and Buckminsterfullerenes are defined by symmetry as they exist in various constitutions[40–43], mainly differentiated by symmetry[1,41–44]. For these carbon allotropes the differences are pronounced[45], but for all organic molecules that are identified using NMR, the symmetry of the molecular structure must be known to interpret the spectrum.

From the earliest models of the electronic structure of transition metal complexes and metal clusters symmetry arguments have been used[46–48], also the metal centers in metalloproteins are differentiated by symmetry[49]. Where crystal field theory fully establishes the theoretical basis for symmetry in the description of coordinating atoms and the electronic structure[50–52], the molecular structure is not as easily defined. Jahn-Teller effects reduces otherwise 'perfect symmetry', and it has been reported repeatedly that tools like the Continuous Shape Measure (CShM) do not work in these cases[53]. That the symmetry of the transition metal compounds is important is clear in magnetic materials and from measurements using e.g., EPR and Mössbauer spectroscopy[11,54–59]. Even high energy X-ray spectroscopies rely on symmetry considerations when modeling the experimental results[60,61].

Magnetic properties are particularly sensitive to symmetry. Thus, symmetry is particularly important in the area of single molecule magnets (SMMs)[62–67]. SMM performance, in particular blocking temperature, has been proposed to depend on symmetry[68,69]. But it remains unclear if it is SMM linearity or if it is other symmetry operations that are most important, a question that is likely to vary depending on the central metal[70,71]. As quasi-symmetry often is the only measure reported[72,73], it is hard to determine which symmetry is best. Particularly, in cases where the distinction between coordination geometry and molecular geometry is not clear[68,74–76]. The area of SMMs needs accurate determination of symmetry and it is clearly stated that CShM, e.g., as implementation in the SHAPE software, is not good enough[69,77].

The electronic structure of lanthanide(III) compounds is very sensitive to small changes in structure and symmetry. This is apparent in the lanthanide(III) based SMMs[62,64–69], but has also been shown in fundamental studies of ytterbium(III) Yb[78–81] europium(III) Eu[82,83], and neodymium(III) Nd[84,85]. In all cases, the determined properties are related to the reported molecular structures.

## The molecular entity

IUPAC defines constitution and conformations, which in turn defines all molecules and materials through the specific connectivity of distinct sets of atoms placed at specific relative coordinates (x,y,z). When considering structure, we can ignore bonding and all other interatomic interactions, as these becomes a consequence of the set of atoms and atomic positions.

IUPAC further defines a molecular entity, see Fig. 1 which is a molecule or material of a specific constitution that is distinctly different from another molecule or material of the same constitution. That is two molecular entities are the same groups of atoms, but where the electronic structure or the atomic positions give rise to two sets of properties **Ø** that we can discriminate in a measurement.

The atomic positions (x,y,z) we define as the molecular structure $\mathbf{Q}$(x′,y′,z′), when the atomic positions are aligned to the main symmetry axis (z′). Each molecular entity corresponds to a set of molecular structures $\mathbf{Q}_N$, sets of atomic positions that all give rise to the same measured property **Ø**. To define the molecular entities, we need to be able to define, orient and compare the corresponding groups of atomic positions, and we define the molecular structure **Q** to do this comparison.

Two molecules—methane and chloroform—with different constitutions are by definition different molecular entities. Compounds with identical constitution can be different molecular entities e.g., molecular entity **1**, with the molecular structure set $\mathbf{Q}_1$, has a property **Ø** that is different from molecular entity **2**, with the molecular structure set $\mathbf{Q}_2$ and the property $\mathbf{Ø}_1 \neq \mathbf{Ø}_2$. We know we have two molecular entities as we have measured two sets of properties. To compare molecular entity **1** and **2** their molecular structures $\mathbf{Q}_1$ and $\mathbf{Q}_2$ must be known and in the same coordinate system. If we measure three properties, there will be three molecular entities etc.

By definition a set of molecular structures $\mathbf{Q}_N$ describes a single molecular entity as long as it gives rise to the same distinct measurable property **Ø**. None of the tasks involved in defining a molecular entity is trivial. The molecular structures must be known, and the measurement can be conditional. In an ensemble of molecules, the variation in molecular structures may be so small that only one property is measured e.g., the boiling point of water. While another measurement reveals that several molecular entities are present e.g., the IR spectrum of bulk water[26,27]. With properties and atomic positions in hand, these must be compared in the correct coordinate system.

Examples of compounds that are different molecular entities include: *i*) benzene in the ground state $S_0 \neq$ benzene in the first excited singlet state $S_1 \neq$ and benzene in the triplet state $T_1$. *ii*) water ice-III ≠ water ice-IV. *iii*) neodymium(III) nonaaqua ion in the tricapped trigonal TTP form ≠ neodymium(III) nonaaqua ion in capped square antiprismatic cSAP form. And *iv*) a transition metal complex with identical connectivity and $O_h \neq$ a transition complex with identical connectivity and $T_d$ symmetry.

## Determining structure

Detailed molecular structure determination is challenging, which we exemplify with methane. Figure 1b shows a methane molecule, a small set of atomic coordinates describing the placement of four protons around a central carbon atom in the coordinate system defined in Fig. 1a. The shape of the molecule is described as a perfect tetrahedron, and by aligning the atomic coordinates in the coordinate system described by the $T_d$ point group, we can show that methane indeed has $T_d$ symmetry. As shown in Fig. 1b, the molecular structure **Q** of methane is reproduced by all the symmetry operations of the $T_d$ point group. Thus, we can describe methane as a single molecular entity with $T_d$ symmetry. Note that methane excited to vibrational states that

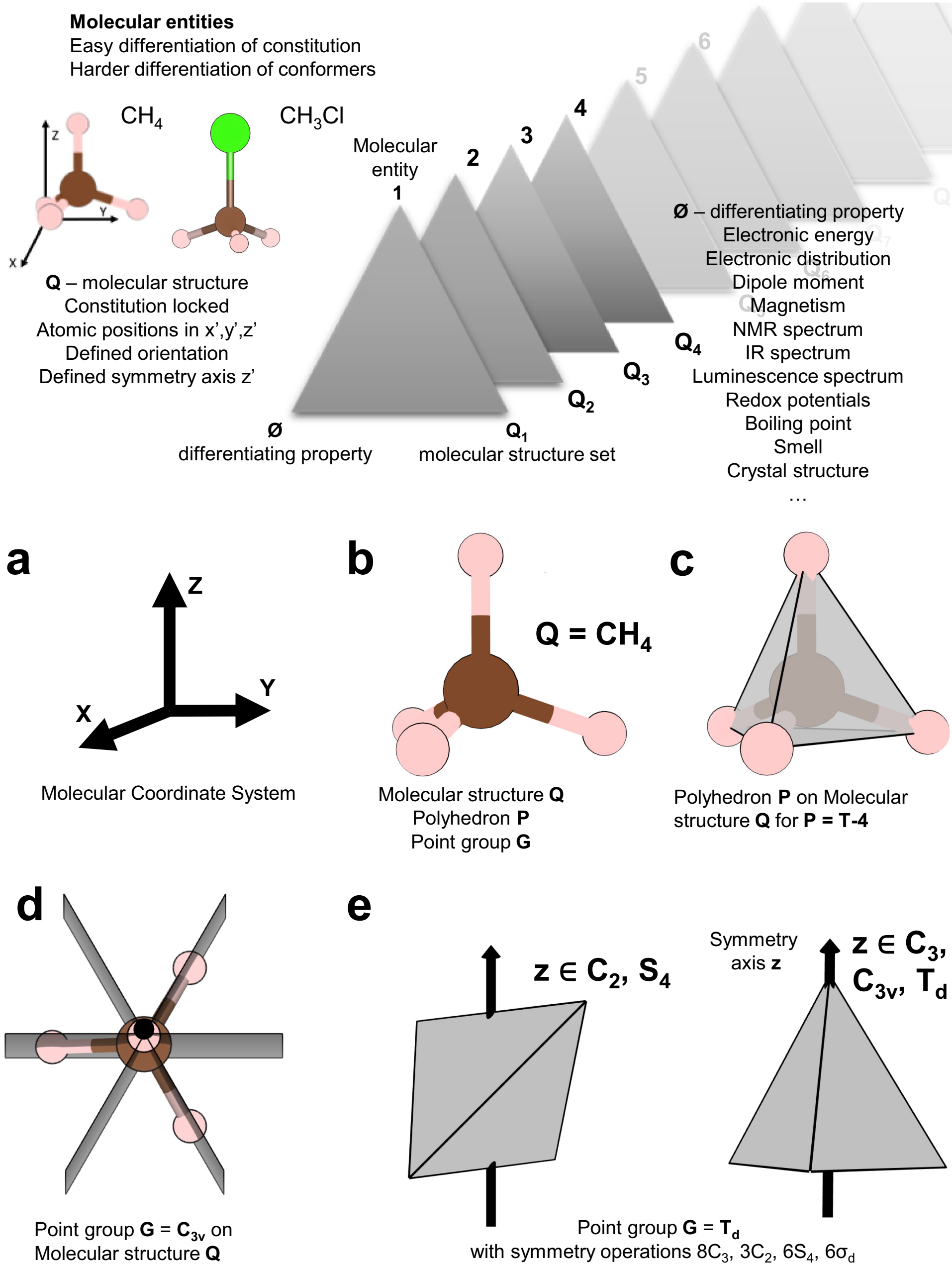


**Fig. 1 | The definition of the molecular entity using the molecular structure Q with the experimental measurable Ø, and the steps to determine coordination geometry P and molecular symmetry i.e., point group G of methane.** **a** Defining the molecular coordinate system (x,y,z). **b** Identifying the molecular structure **Q**. **c** The structure **Q** can be determined to be identical to a polyhedron **P** = T-4 (a tetrahedron). **d** The structure **Q** is highly symmetric with symmetry elements like: a $C_3$-rotational axis and three mirror planes. **e** The symmetry of methane is described by **G** = $T_d$ and **Q** is reproduced by all 24 symmetry operations in **G** = $T_d$. Note that the symmetry axis **z** depends on the symmetry operation, while the molecular Z axis remain constant. Atom legend. carbon = brown, hydrogen = pink, chlorine = green.

break this symmetry must be considered as separate molecular entities, as e.g., the symmetry and IR spectrum will be different for vibrationally excited methane.

The atomic coordinates of methane in the correct coordinate system can, using symmetry operations, be shown to have $T_d$ symmetry. All other relevant molecular structures of methane will be of higher complexity. Therefore, we must define a single measure of symmetry and start reporting on molecular structures using this measure. Further, to make the measure relevant, we must often assume that we are working on static molecular structures. The alternative is to drown in complexity e.g., if we are to differentiate between the many different thermally excited methane molecules in a sample.

Beyond ideal molecules with symmetry (e.g., methane with $T_d$ symmetry), the complexity in determining structure increases. Even with the atomic coordinates of a molecular structure in hand, describing the molecular structure is not trivial, and several approaches are in use. Often, the structure is simply determined ad hoc or relies exclusively on the crystal structures and crystal symmetry. If symmetry is not present, the term pseudo-symmetry is used to assign non-perfect symmetry to molecules and crystals. And even though symmetry is fundamentally a binary concept—it is either present or absent, experiments indicate that approaching a given symmetry is enough for symmetry to be defining the observed properties[9,10,12,17,18,86–89]. Thus, moving beyond absolute symmetry and defining a quantity that report on the distortion from symmetry in a continuous matter makes sense. That is, determining symmetry as a continuous property has merit.

The methods available to quantify symmetry include the Continuous Symmetry Measure (CSM)[90], which aligns the atomic coordinates and determines distortion from symmetry. While CSM in principle is exact across all symmetries for all structures and mathematically elegant, it has limitations arising from how it is defined and the most recent implementation is limited to cyclic symmetry point groups[91]. An alternative to a symmetry method is the Continuous Shape Measure (CShM) that measure how close a molecular structure **Q** is to a set of selected structures **P'**. The CShM tool has severe limitations, as the current implementations are restricted to the inner coordination sphere and does not take the nature of the coordinating atoms into account. Thus, we prefer to quantify symmetry with the Continuous Symmetry operation Measure (CSoM). The CSoM quantifies the symmetry of any molecular structure, and we propose that CSoM determined symmetry and that the CSoM determined symmetry deviation $\sigma_{sym}(\mathbf{Q},\mathbf{G})$ is the best descriptor of both structure and molecular properties.

To measure symmetry with CSoM, a molecular structure **Q** is defined as a set of atomic positions. The CSoM software aligns the molecular structure **Q** in the appropriate molecular coordinate system for all point groups, and reports which symmetry i.e. which point group **G**, that is the most relevant for this molecular structure. Further, the CSoM approach then provides a numerical deviation from symmetry—a $\sigma_{sym}(\mathbf{Q},\mathbf{G})$-value, for all selected point groups. The $\sigma_{sym}(\mathbf{Q},\mathbf{G})$-value defines how well the molecular structure **Q** is described by each **G**. Thus, the CSoM places the molecular structure **Q** on a continuum of $\sigma_{sym}(\mathbf{Q},\mathbf{G})$-values approaching symmetry which is identified as $\sigma_{sym}(\mathbf{Q},\mathbf{G}) = 0$. By using the CSoM software, we determine and define the symmetry, constitution, and confirmation of each molecular entity in a consistent framework on an absolute scale.

### Practical considerations when comparing molecular structures

The structure of compounds can come from many sources and arise from methods with or without restraints. The most important prerequisite, when determining the molecular structure of a molecular entity, is that the crystal structure was not solved—and the computed coordinates were not generated—with any restrictive symmetry. This may lead to artifacts, as these methods per definition generate atomic coordinates of ideal symmetry.

### Constitution – how large is the molecular structure?

The same molecular entities cannot have different constitutions. Thus, it is a requirement that molecular structure comparisons take the constitution into account. Nevertheless, it can be relevant to compare molecular structures with different central ions e.g., in the *d*-block and the *f*-block. Historically, comparison of metal complexes that differ in donor atoms has been made, and comparisons have been done based on coordination polyhedra alone. These simplifications may be relevant in a descriptive context, but not for describing molecular entities. Here, the question becomes how much of a compound is needed to describe the molecular structure responsible for the observed property. Is the molecular entity defined by the local structure around a metal ion (the first coordination sphere), is the second coordination sphere needed, or should the full complex and solvent be included? Similarly, for solids, is the complex enough, is the symmetry equivalent, should the full unit cell or multiple unit cells be used, and is the solvent in the structure important? The questions are open, as we must determine the size of the molecular structure that gives rise to a specific property, but the method for comparing molecular structures should be able to handle the different types of input.

Note that it is assumed that all properties and structures under consideration are known, isolated and arise from distinct species[92,93].

### Scaling

Comparing atomic coordinates based on reference structures i.e., using CShM, inherently requires some form of scaling. The same applies if molecular structures with different elements are compared using position or distances alone. While scaling itself is not problematic, it is an additional parameter that must be declared and does create an opportunity for errors to occur. Comparing symmetry, i.e., using CSM and CSoM, does provide a normalized measure but the comparisons themselves do not require scaling.

### The yardstick

The true value of molecular structure determination come at scale. Defining a common yardstick that is universal and readily applied will allow a field to start comparing their important properties to e.g., symmetry on a common scale. The key property of the yardstick is that it is agnostic and translatable. Our recommendation is that CSoM should be widely used at the expense of the other approaches. The reason why should become apparent below.

## Methods used to determine structure

### The continuous symmetry measure: CSM

CSM is a direct measure of symmetry, as defined by a measure for the shortest distance the atomic coordinates have to move to construct a polyhedron with evaluated symmetry. In technical terms: CSM measures the deviation to a point group **G** defined as the Euclidean distance between the coordinates of the evaluated structure **Q** and the coordinates of the best aligned G-symmetric polyhedron. The deviation S(**Q**,**G**) is defined by the average squared distance between each vertex k of **Q** and **P** best aligned to each other with respect to $|Q_k - P_k|$[27]. The measure is defined in Eq. 1, and the average of Euclidean distances is a normalized root mean square value, where 0 is perfect symmetry.

$$S(\boldsymbol{Q},\boldsymbol{G}) = \min_P \sum_{k=1}^{N} \frac{|Q_k - P_k|^2}{|Q_k - Q_0|^2} \times \frac{100}{N} \, where \, P|_G \tag{1}$$

**Q** is a sample structure with k vertexes, **G** is the point group symmetry to which the structure is evaluated to, **P** is a structure with k vertexes that is restricted to the symmetry $\boldsymbol{P}|_{\boldsymbol{G}}$ with minimal distance to **Q**, and N is the number of vertexes k in the polyhedron. The definition is based on early work developed in 1992 by D. Avnir et al[90,94].

Illustrations of the CSM calculations are shown for $CH_4$ and $CH_3Cl$ in Fig. 2a. $CH_4$ is perfectly $T_d$ symmetric and the CSM value therefore yields $S(CH_4,T_d) = 0$. The tetrahedron is the only $T_d$ symmetric polyhedron with four vertexes. The CSM value for $T_d$ symmetry in Coordination Number (CN) = 4 complexes are therefore defined with the tetrahedron T-4. $CH_3Cl$ is $C_{3v}$ symmetric and does not contain $T_d$ symmetry and the CSM value for $T_d$ is therefore required to be nonzero. As the tetrahedron is the only $T_d$ symmetric four vertex polyhedron the CSM value is again obtained by evaluating the CHCl with T-4. The obtained value is $S(CH_3Cl,T_d) = 7$ as shown in Fig. 2a. The structure is found to be even worse described by $D_{4h}$ symmetry: $S(CH_3Cl,D_{4h}) = 35$. For $D_{4h}$ symmetry the polyhedron used is the square, which is the only $D_{4h}$ symmetric polyhedron with four vertexes. The actual symmetry of $CH_3Cl$ is $C_{3v}$, but this is not readily found with CSM. The triangular pyramid vTBPY-4 is a $C_{3v}$ symmetric polyhedron, however, an infinite span of different vTBPY-4 polyhedra exists—all with $C_{3v}$ point group symmetry. To find the one with the minimal distortion, to satisfy Eq. (1), different search algorithms can be employed.

The original approach that creates **P** from **Q** with respect to a particular symmetry used the so-called 'folding-unfolding' algorithm [86,90,95]. For four vertexes this works well and the $C_{3v}$ symmetry can readily be assigned for $CH_3Cl$ using this method. Using the correct vTBPY-4 polyhedron the structure is indeed found to be

https://doi.org/10.1038/s41467-025-66054-7

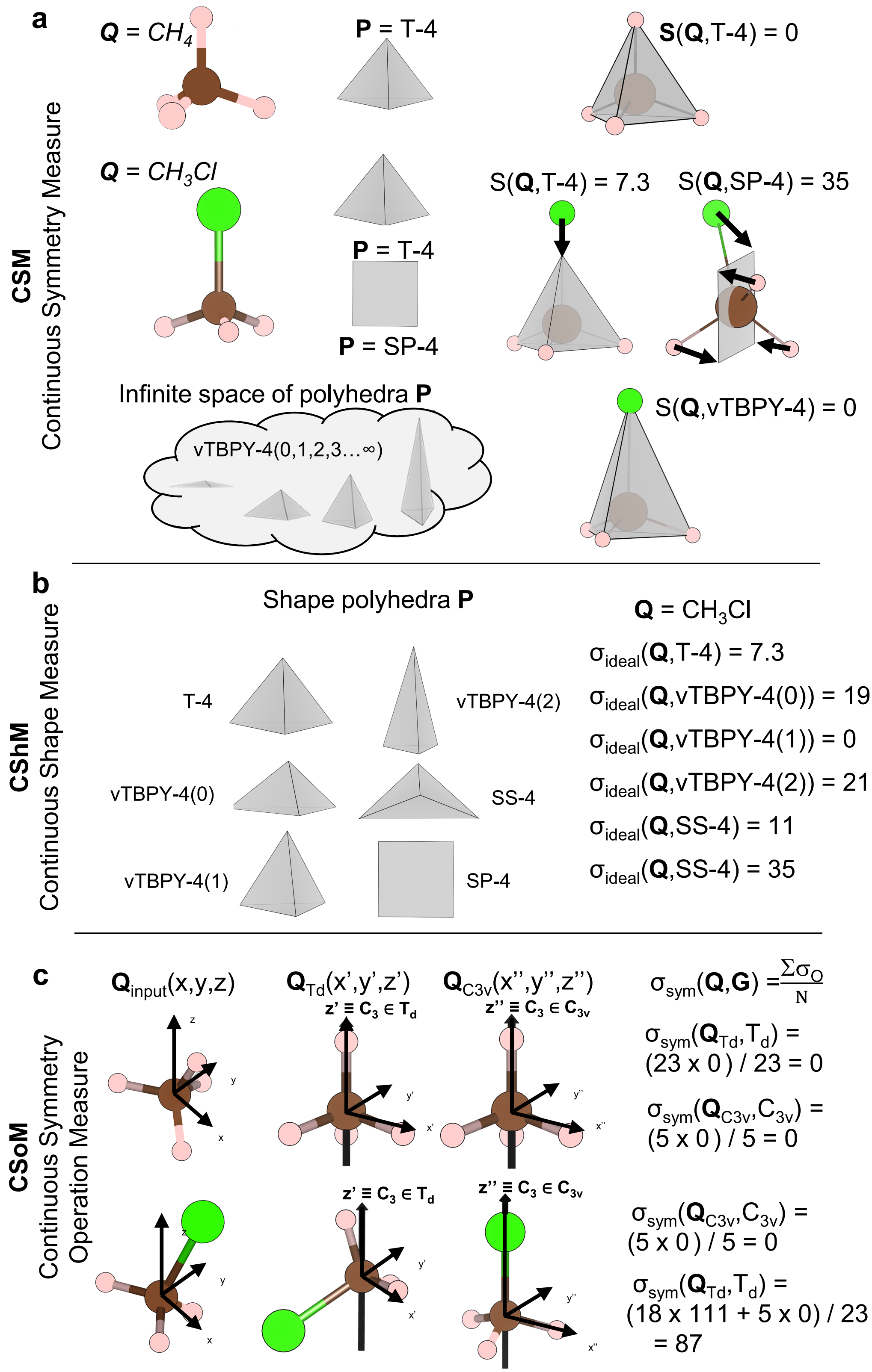


**Fig. 2 | Three approached for determining and describing molecular structure.** Determining molecular structure **Q** and symmetry point group **G** of methane and chloroform either using polyhedral **P** with the (**a**). CSM, **b** CShM or symmetry with **c** CSoM. CSM provides S-values, CShM provides σideal-values and CSoM provides σsym-values as deviations from ideal fit to **P** or the symmetry **G**. Atom legend: carbon = brown, hydrogen = pink, chlorine = green.

$C_{3v}$-symmetric: S($CH_3Cl$, $C_{3v}$) = 0. Beyond four vertexes, the algorithm becomes computationally heavy. These issues can be overcome with intelligent algorithmic design. A recent tool, ChemEnv[96] by Waroquiers et al., drastically reduces the amount of permutations by separating groups of atoms within the chemical environment. The most recent implementation of CSM[91] by Tuvi-Arad et al. scans only chemically relevant permutations that maintain the connectivity of the molecule allowing the analysis of larger molecules, but is limited to cyclic symmetry point groups. CSM as implemented by Tuvi-Arad is used below.

### The continuous shape measure: CShM

An alternative to the CSM is the CShM. CShM is not a measure of symmetry, but simply a geometrical deviation value to a specific reference polyhedron. A search for the perfect polyhedron is therefore not performed. Instead, a manually selected reference polyhedron **P** is used to compare to the input structure **Q**. The equation to calculate the geometrical deviation $\sigma_{ideal}$(**Q**,**P**) is defined in Eq. 2 and is mathematically identical to CSM. The difference is that no minimizations or searches are performed for **P** prior to the evaluation, and the deviation value is therefore not to a symmetry group. Instead, it is to a specific polyhedron, which may be symmetric, but is not restricted to be symmetric[94,97,98].

$$\sigma_{ideal}(\boldsymbol{Q},\boldsymbol{P}) = \sum_{k=1}^{N} \frac{|Q_k - P_k|^2}{|Q_k - Q_0|^2} \times \frac{100}{N} \quad (2)$$

The CShM has been implemented in many ways[6,86,90,94,95,98–111], but the most commonly used algorithm is the one provided by the Shape v2.1 software[103]. The algorithm is easy to use and the measure has proven to be able to quantify the approximate structure of complexes with coordination numbers from 4 to 10 [99,100,106,112–116]. Fig. 2b illustrates the calculations for $CH_3Cl$ which is evaluated with six different four-vertexes-polyhedra: T-4, vTBPY-4(0), vTBPY-4(1), vTBPY-4(2), SS-4, and SP-4. As was found with CSM, the $CH_3Cl$ structure matches perfectly with a specific tetrahedron vTBPY-4. However, an infinite span of vTBPY-4 polyhedra exists with an infinite set of ratios between the distances between the apex vertex to the vertexes in the base. Figure 2b displays four different of such polyhedra, as indeed, the T-4 is also a special case of the infinite span of vTBPY-4. Therefore, finding the ideal reference polyhedron is not trivial and as the reference polyhedra are manually selected, no guarantee to select the correct polyhedron can be given.

The use of CShM and the SHAPE software should therefore be performed with caution. Three different vTBPY-4 polyhedra, excluding T-4, are used as reference polyhedra in Fig. 2b. The best reference is obviously vTBPY-4(1) resulting in $\sigma_{ideal}$($CH_3Cl$, vTBPY-4(1)) = 0 and $CH_3Cl$ can be perfectly described as vTBPY-4, but without a search procedure any value could in principle be obtained. When the best polyhedron is found the special case where CShM and CSM coincide is obtained. In this case the vTBPY-4(1) geometry satisfy Eq. 1 and the CShM is in this specific case identical to CSM evaluated with C3v symmetry for the $CH_3Cl$ structure Fig. 2b displays three cases where CSM values coincide with CShM values: T-4 is the CSM with $T_d$ symmetry as this is the only 4-vertex polyhedron with $T_d$ symmetry, SS-4 with $D_{4h}$ symmetry as this is the only 4-vertex polyhedron with $D_{4h}$ symmetry, and vTBPY-4 with $C_{3v}$ symmetry as this is the best 4-vertex polyhedron to describe the structure with $C_{3v}$ symmetry. A recent development solves this issue. A software called Polynator uses flexible polyhedral models to determine the coordination geometry[117]. The polyhedra are stretched and bent within the restrictions enforced by the definition of the polyhedron, until a best fit model has been determined. Polynator thus solves the issue of finding the best vTBPY-4 polyhedron in Fig. 2b. However, the implementation employs a different yardstick than CSM, CShM, and CSoM, and the numeric results are not directly comparable to these methods.

The CShM is much more practical to use than CSM, but caution should be exerted when the calculated deviation values are interpreted. Typically, the calculations are arbitrarily selected sets of polyhedra and may not accurately reflect the distortion from symmetry. Another issue with CShM is the inability to evaluate structures to symmetries for which a polyhedron does not exist or for which so many exist that selecting the correct one is impractical. A list of all polyhedron, as provided in Shape v2.1[103] from CN = 2 to 12 is provided as Supplementary Information Table S1. Here multiple relevant examples can be considered: For CN = 8 the $D_{4d}$ symmetry contains only a single structure and no good polyhedron with simple $C_4$ symmetry can be constructed and evaluated. The $D_{2d}$ symmetry has two different polyhedra: TDD-8 and JGBF-8. These two are very different in shape and the measure therefore potentially gives very different values. At CN = 9, four $C_{4v}$ polyhedron exists and two $D_{3h}$, but no $C_3$, $C_4$, and $D_3$ symmetric polyhedra exists. In general, no polyhedron with simple rotational symmetry exists: i.e. the $C_2$, $C_3$, …, $C_n$ symmetries cannot be identified. To measure approximate symmetry, CShM often provides an incomplete description. This issue increases for higher coordination numbers. It is, as the name indicates, a measure of shape and not symmetry. An alternative approach is to characterize the shape of the first coordination sphere with ellipsoids irrespective of coordination number and geometry[118]. The method is reported to be especially useful for octahedral distortions such as tilt and strain.

### CShM: orientation of input structure Q

In order to orient the input structure **Q** with respect to the reference polyhedron, it is necessary to implement a series of algorithms. To compare **Q** and **P** a common molecular coordinate system (x,y,z) must be defined and a normalization must be implemented. The molecular coordinate system is defined from an origin (0,0,0), which can be manually selected, defined by the central atom, or by the center of mass. The size of **Q** must then be normalized to the polyhedron **P**. This can be accomplished in two ways. The first method is to normalize the average of the coordinates to have a length of 1 from the origin. The second method is to define all coordinates to be on the unit circle. The correct choice may vary between systems. When **Q** is aligned and normalized, it must be rotated to reach optimal alignment with **P**. This alignment is achieved by minimizing the rotation matrix between the two sets of coordinates, e.g. with the Kabsch algorithm[119].

To minimize correctly, the labels for each coordinate set in **Q** must remain valid after each step. Thus, each rotation must occur while tracking the labels. While minimizing the rotation matrix between **P** and **Q** is a simple task in itself, tracking labels for each coordinate carries a computational that grows with N!, where N is the number of coordinate sets/labels. For a CN = 10, N! approach 4 million, which makes the approach impractical. Orientation of the input structure of CShM is described in detail elsewhere[94].

### The symmetry operation measure: SoM

SoM does not measure the difference to a different polyhedron and is fundamentally different to CSM and CShM. SoM is a measure of how well a symmetry operation transforms a structure back into itself[120]. The mathematical formula is identical to CShM as seen in Eq. 3. The important difference is the reference polyhedron, which is replaced by the symmetry operated structure $\hat{O}_S$**Q**, where $\hat{O}_S$ is a symmetry operation and **Q** is the structure. The measure is thus a descriptor of how good a specific symmetry operation can be used to describe the structure.

$$\sigma_O\left(\boldsymbol{Q}, \hat{O}_S\boldsymbol{Q}\right) = \sum_{k=1}^{N} \frac{\left|Q_k - \hat{O}_S Q_k\right|^2}{|Q_k - Q_0|^2} \times \frac{100}{N} \quad (3)$$

Unlike CShM, SoM accounts directly for symmetry as the measure evaluates how well a structure is reproduced by a symmetry operation. The method has seen various algorithmic developments[3–5,8,121], and has been used to evaluate the symmetry of simple structural coordinates[1,7,108] and has even been generalized to wavefunctions[6,44,122].

The issue with SoM is the limitation to only a specific symmetry operation and is therefore not a measure to a complete point group. More importantly, knowledge of the principal axis and orientation of the structure is needed to use SoM. Evaluation with rotational symmetry only really makes sense if the symmetry operation is applied along the best possible symmetry axis.

### The continuous symmetry operation measure: CSoM

CSoM is a development on the SoM. It is defined as an average of all SoM for symmetry operations in an entire point group, excluding the identity. The CSoM, $\sigma_{sym}(\mathbf{Q},\mathbf{G})$, is defined in Eq. 4.

$$\sigma_{sym}(\boldsymbol{Q},\boldsymbol{G}) = \sum_{s=1}^{N} \frac{\sigma_O\left(Q, \hat{O}_S Q\right)}{N} \tag{4}$$

$\sigma_{sym}(\mathbf{Q},\mathbf{G})$ is a direct measure of approximate G-symmetry of the structure Q. $\sigma_{sym}(\mathbf{Q},\mathbf{G})$ does not need a reference structure, but an input of a selected point group. Note, that as the individual SoM values for all operations within the selected point group are averaged, the individual contribution of each symmetry operation weighs less in the total $\sigma_{sym}(\mathbf{Q},\mathbf{G})$ value for larger point groups.

Figure 2c shows the CSoM results of $T_d$ and $C_{3v}$ symmetry on the $CH_4$ and $CH_3Cl$ molecules. To evaluate the structure, they must first be aligned with the principal axis z of the point group symmetries. Both $T_d$ and $C_{3v}$ can share a $C_3$ axis, such that the z axis may coincide or be different between point groups.

For $CH_4$, all 24 and 6 symmetry operations of the $T_d$ and $C_{3v}$ point groups respectively give a SoM result, a $\sigma_O$-value, of zero. Therefore, the CSoM methods gives a $\sigma_{sym}$-value of zero for these two point groups that is $\sigma_{sym}(CH_4,T_d) = (23 \times 0) / 23 = 0$ and $\sigma_{sym}(CH_4,C_{3v}) = (5 \times 0) / 5 = 0$. $CH_4$ is therefore found to contain both $C_{3v}$ and $T_d$ symmetry.

$CH_3Cl$, however, is not $T_d$-symmetric and the $\sigma_{sym}$-value for $T_d$ is nonzero. Figure 2c shows the structure aligned in the best possible orientation with respect to the operations in $T_d$. Most of the individual SoM values for $CH_3Cl$ evaluated with $T_d$, give numerical values of 111, which is a very large numerical value that reflects poorly reproduced structures. However, some of the individual operations (i.e., two $C_3$ and three $\sigma_d$ operations) are found to perfectly reproduce the structure with a SoM value of $\sigma_O = 0$. The average deviation, which is the CSoM value gives $\sigma_{sym}(CH_3Cl,T_d) = (18 \times 111 + 5 \times 0) / 5 = 87$.

The structure does not have $T_d$ symmetry. Moving to $C_{3v}$ symmetry the best principal axis is through the $C_3$ rotation of the C-Cl bond. It is found that $\sigma_{sym}(CH_3Cl,C_{3v}) = (5 \times 0) / 5 = 0$ and $CH_3Cl$ is found to have perfect $C_{3v}$ symmetry. CSoM is thus a direct measure to evaluate symmetry as compared to CShM and CSM, which are only indirect symmetry measures. They rely on reference structures with certain symmetries.

### CSoM: orientation of input structure Q

The biggest issue with CSoM is finding the best orientation of the coordinates to best match the point group symmetry. To address this issue, a minimizing algorithm to find the best possible principal axis was developed. The minimization algorithm is described in detail in ref. 2.

### Other symmetry measures

Numerous methodologies exist that determine distortions from ideal shape and symmetry. The recently developed PorphyStruct is a tool for the analysis of specifically non-planar distortion modes in porphyrinoids[123]. Using a normal-coordinate structure decomposition technique it is able to characterize the distortions in these $D_{4d}$ symmetric molecules arising from vibrational modes. A similar approach is found for the program Symmetry-Coordinate Structural Decomposition, which is able to measure distortions of a irreducible representation of a point group[124]. These techniques may allow symmetry measures to go beyond static structures.

Alternative methods are in use that quantify the similarity of multiply constitutionally identical molecules within the same unit cell (i.e. Z′ > 1 crystals). Identifying pseudo-symmetries within the unit cell of Z′ > 1 crystals can be done through different methodologies[125–128]. One such example is the computer program CRYSTALS[125], which match two sets of coordinates within the same unit cell. The measure is similar to CShM, with the reference structure that is a same-constitution structure within the unit-cell.

Finally, Hyperspace Recognition is used to compare the similarity in structure between molecules of different constitution[129]. This method goes beyond the measure of simple geometric symmetry, but is not intrinsic to the specific molecular entities.

## Results

Recommending the use of CSoM as a general measure of molecular structure and as a tool to identify molecular entities requires that we 1) explore the limits of the other measures of molecular structure that are in use, and 2) attempt to quantify which numerical values of $\sigma_{sym}$ that represent the presence/absence of symmetry. We must remember that symmetry is a binary measure, however, molecular properties can show symmetry even though the geometrical symmetry of the molecular entity is imperfect.

We start by exploring water. The water molecule, data from molecular dynamics simulation of liquid water, and several forms of water ice. We remain in the *p*-block and demonstrate that CSoM can be used on benzene and larger polycyclic aromatic hydrocarbons (PAHs), which is not possible using CShM. Then, we move to the *d*-block looking at coordination polyhedra and molecular structures of transition metals, before concluding in the *f*-block. Using inorganic nomenclature to describe the number of atoms around the central atoms, we move from coordination numbers (CN) 2, 3 and 4 in water and PAHs over to CN 4-8 in the *d*-block to CN 5-12 in the *f*-block thus demonstrating that the CSoM method can be used to describe and define all molecular structures using symmetry. The molecules investigated are shown in Fig. 3.

### Water – $H_2O$

Water is a $C_{2v}$ symmetric molecule. The ideal $H_2O$ molecule is shown in Fig. 4a. This is the a priori result and the result obtained if the structure is opposed in silico with e.g., the UFF force field. We used UFF implemented in Avogadro using steepest descent to generate the in silico structure. The CSoM value for this optimized water structure is 1E-10, which is the numerical cut-off value of the minimization procedure in the CSoM implementation and is therefore regarded as 0. $H_2O$ has three vibrational degrees of freedom. An asymmetric stretch, a symmetric stretch, and a symmetric bend. Only the asymmetric stretch distorts the structure from $C_{2v}$ symmetry.

To explore how much the molecule naturally deviates from $C_{2v}$, we sampled the structures from a molecular dynamics simulation[34]. The simulation is of 266,667 water molecules described with force fields (SPC) with a periodic boundary condition and with a density of 1 kg/L and an average temperature of 300 K. Calculating the CSoM value of a subset of 2000 water molecules from this simulation we found an average of $\sigma_{sym}(\text{water},C_{2v}) = 1.3\text{E-}05$, see Fig. 4b. While these are orders of magnitude larger than the optimized geometry they are still incredibly small, and we conclude that water maintains near-perfect $C_{2v}$ symmetry in silico. It is worth noting that if the mirror plane is evaluated with CSoM, $\sigma_{sym}(\text{water},\sigma) = 1\text{E-}10$ is found. As three points always span a plane, no distortions from ideal structure can remove

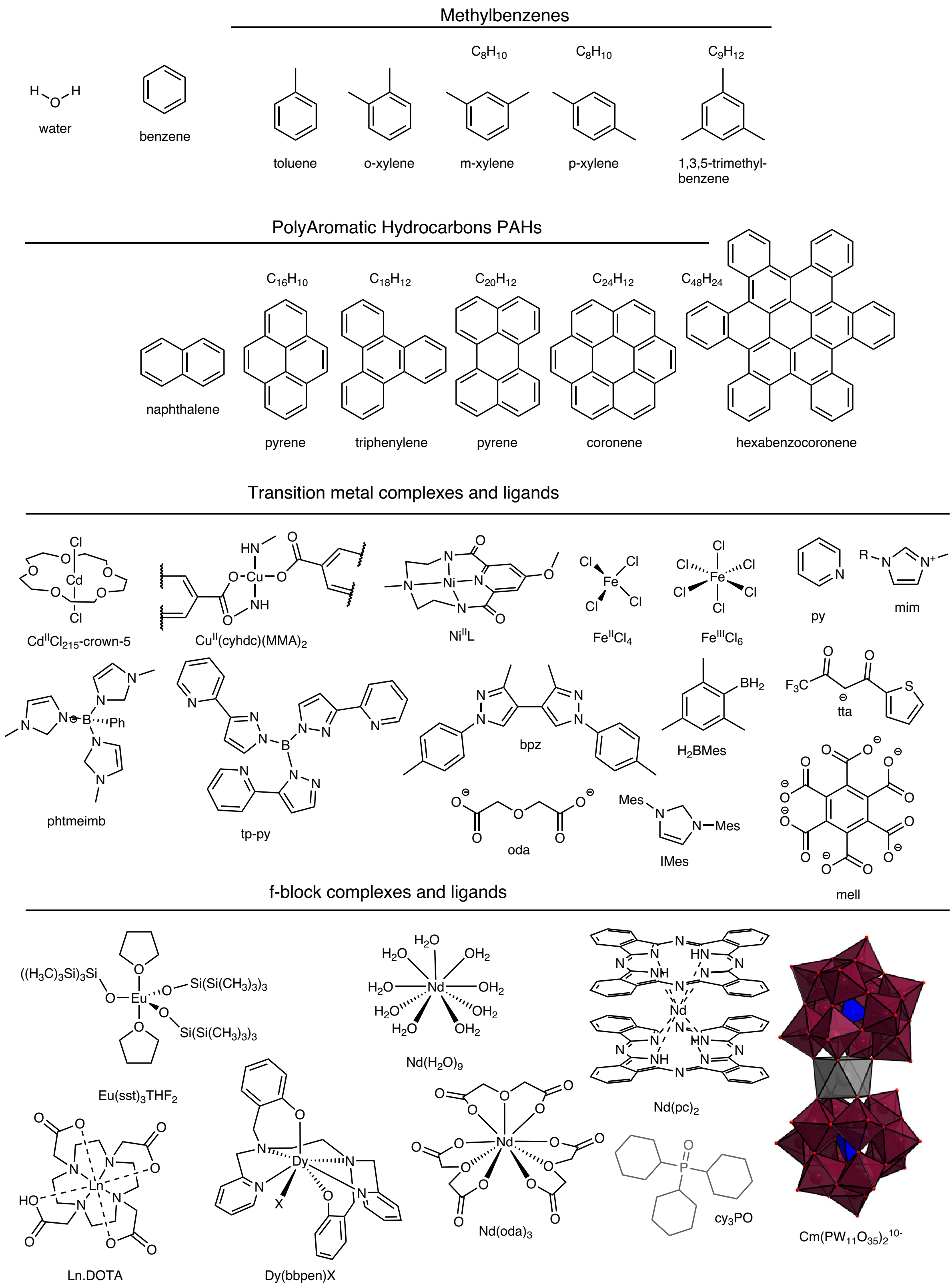


**Fig. 3 | Molecules and complexes investigated with the CSM, CShM, and CSoM method.** Atom legend: oxygen = red phosphorous = blue, curium coordination polyhedra = gray, tungsten coordination polyhedra = deep purple.

the $C_\sigma$ symmetry. Furthermore, the orthogonal plane to the plane that the three points spans will be symmetrically equal to the $C_2$ operation. Thus, the $C_{2v}$ point group of triatomic molecules is always related to the $C_2$ point group with $\sigma_{sym}(\mathbf{Q},C_{2v}) = 2/3\sigma_{sym}(\mathbf{Q},C_2)$.

The analysis of water also shows how CSoM can quantify minute deviations from perfect symmetry that arise in "real and perfect" systems, see how the structure of the 2000 water molecules in the MD simulation is readily analysed using CSoM.

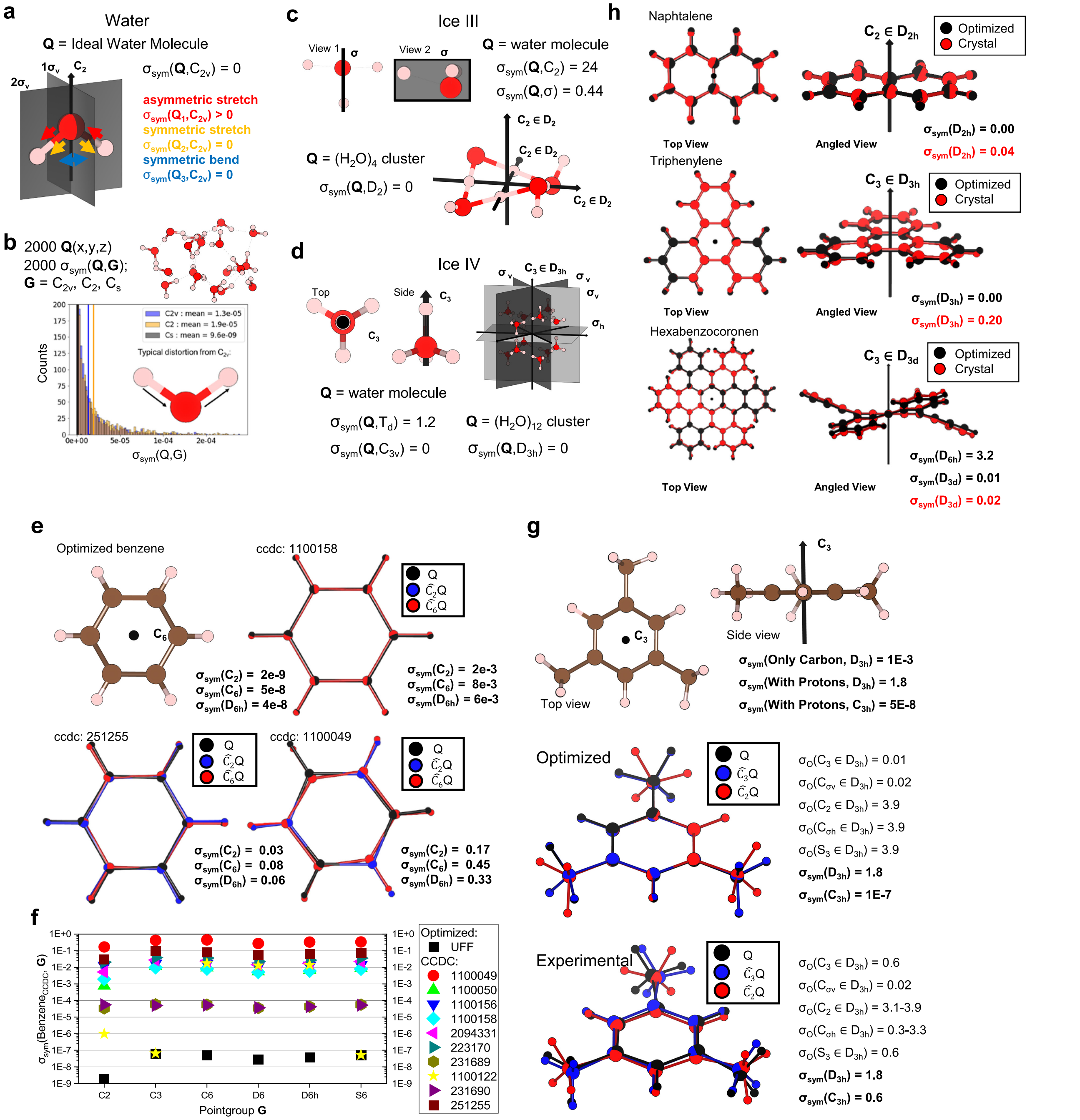


**Fig. 4 | Water and aromatic molecules evaluated with CSoM for point group G providing deviations from the symmetry as σsym-values. a** Ideal geometry of $H_2O$ with $C_{2v}$ symmetry that can only be broken by the asymmetric stretch. **b** 2000 $H_2O$ molecules from a molecular dynamic simulation[34] evaluated with the $C_2$, $C_\sigma$, and $C_{2v}$ point groups, illustrating the small effect of the asymmetric stretch on symmetry. **c** The symmetry of water molecules and clusters in ICE III and **d** the symmetry of water molecules and clusters in ICE IV[130,131]. **e** CSoM analysis of a single in silico optimized benzene as well as three crystal structures showing the visual numerical CSoM output. **f** CSoM analysis of 11 benzene molecular structures[172–174,192–195]. **g** CSoM output for the in silico and crystal structure of 1,3,5-trimethyl-benzene. **h** CSoM analysis for in silico optimized structures and experimental crystal structures of naphthalene, triphenylene, and hexabenzocoronene. The symmetry aligned structures from in silico and the crystal structures are overlaid for visual comparison[132,133,180] Atom legend: carbon = brown, hydrogen = pink, oxygen = red.

## Water ice – $(H_2O)_n$

While water is a $C_{2v}$ symmetric molecule in the gas phase and in silico, it does not retain a $C_{2v}$ site symmetry in ice. Figure 4c, d shows the analysis of two different forms of ice, ICE-III and ICE-IV[130,131].

Considering a single water molecule, ICE-III has the oxygen atom bound to three hydrogen atoms maintaining only partial mirror symmetry $\sigma_{sym}(\text{water},\sigma) = 0.44$. Moving beyond the individual molecule, a supramolecular unit with four water molecules has perfect $D_2$ symmetry $\sigma_{sym}(\mathbf{Q},D_2)$ <1E-10 = 0.

In ICE-IV the individual water molecule has an oxygen atom bound to four hydrogen atoms in apparent tetrahedral symmetry, however $\sigma_{sym}(\mathbf{Q},T_d) = 1.2$, which is likely too large to assign tetrahedral symmetry to the individual molecule. Instead, the symmetry was found to be $\sigma_{sym}(\mathbf{Q},C_{3v}) = 0$. Moving beyond the individual molecule, a supramolecular network of 12 $C_{3v}$ symmetric water molecules was found to have $D_{3h}$ symmetry with $\sigma_{sym}(\mathbf{Q},D_{3h}) = 0$. Note that in the crystal structure, each of the four hydrogen positions has an occupancy of ½.

The analysis of ice shows that CSoM works on both individual molecules and larger clusters. In particular it should be noted that CSoM does not require a central atom to work, which allow us to evaluate supramolecular clusters and larger organic molecules.

### Benzene – $C_6H_6$

Benzene is a $D_{6h}$ symmetric molecule. We created an ideal benzene molecule using UFF forcefield in Avogadro using steepest descent. The CSoM analysis found $\sigma_{sym}(D_{6h})$ = 1E-8 which shows that the molecular structure without central atoms can be analyzed, and that UFF produces a $D_{6h}$ symmetric benzene, see Fig. 4e.

Figure 4f shows a CSoM analysis of ten experimentally determined molecular structures of benzene compared to UFF structure. $\sigma_{sym}(\mathbf{Q},C_2)$, $\sigma_{sym}(\mathbf{Q},C_3)$, $\sigma_{sym}(\mathbf{Q},C_6)$, $\sigma_{sym}(\mathbf{Q},D_6)$, $\sigma_{sym}(\mathbf{Q},D_{6h})$, and $\sigma_{sym}(\mathbf{Q},S_6)$ are displayed as these are representative for the differences across the ten structures. In particular $\sigma_{sym}(\mathbf{Q},C_2)$ highlights that the structures are different, while the $C_3$ axis differentiate the ten structures into types, in the same manner as the more symmetric point groups. Note that CCDC 1100122 is best described by $S_6$, which is why this point group was included.

In general, three types of molecular structures were identified. The low symmetry type contains seven structures. A high symmetry type with two structures. And CCDC 1100122 with $S_6$ symmetry.

The molecular structure taken from CCDC 1100049 is the least symmetric in the low symmetry cluster with high CSoM values across all point groups. With a minimum value of $\sigma_{sym}(\mathbf{Q},C_2) = 0.17$, this benzene is not symmetric. The experimental data for CCDC 1100049 has been questioned, and this example may show distortions larger than what is physical for benzene. Nevertheless, the variation between point groups in CCDC 1100049 is mirrored albeit to a lesser extent in all the ten experimentally determined structures, except for the $S_6$ symmetric CCDC 1100122 structure. Figure 4e includes three selected experimental structures. The least symmetric CDCC 1100049, one that is slightly less symmetric CCDC 251255, and the most symmetric of the low symmetry cluster: CCDC1100158. While the distortions are clearly visible in the former two, the latter with $\sigma_{sym}$ < 1E-2 only show minimal differences in the positions of the atoms after the structure has been operated on by the relevant symmetry operators.

### Methylbenzenes – $(CH_3)_n$-$C_6H_{6\text{-}n}$

Having analysed benzene, we increase the complexity slightly with toluene, xylene, and 1,3,5-trimethylbenzene. These molecules contain one, two, and three methyl groups which reduce the symmetry. In particular if the positions of the hydrogen atoms on the methyl groups are considered. Note that the experimental precision in determining the positions of hydrogen atoms vary. Consequently, the accuracy of the structural symmetry becomes limited. However, the following demonstrate that CSoM readily analyse proton positions.

We created idealized molecular structures using UFF forcefield in Avogadro using steepest descent and compared these to experimental crystal structures of the compounds. Table 1 lists the results of the CSoM and CSM analysis. Starting with toluene, the addition of a methyl group breaks the $D_{6h}$ symmetry of benzene. The optimized carbon structure of toluene has $\sigma_{sym}(\mathbf{Q},C_{2v})$ = 2E-2 which indicates perfect $C_{2v}$ symmetry, but the inclusion of protons reduces the symmetry to $C_\sigma$ with $\sigma_{sym}(\mathbf{Q},C_\sigma)$ = 1E-6 and the CSoM value for $C_{2v}$ has significantly increased to $\sigma_{sym}(\mathbf{Q},C_{2v}) = 1.82$. With the addition of the second methylene group, going from toluene to xylene, we have three different cases for *ortho*-, *meta*- and *para*-xylene. Considering first the carbon-only structure and then the full structure with protons, we see the effects of the hydrogen positions. Considering the CSoM analysis for *ortho*-xylene we determine $\sigma_{sym}(o\text{-xylene}^C,C_{2v}) = 0.03$ for the carbon structure, while we find $\sigma_{sym}(o\text{-xylene},C_{2v}) = 0.33$ when including hydrogen atoms. Due to the hydrogen atom the molecular symmetry reduces to $\sigma_{sym}(o\text{-xylene},C_\sigma) = 0.04$. Similar we find that *meta*- and *para*-xylene has $\sigma_{sym}(m\text{-xylene}^C,C_{2v}) = 0.08$ and $\sigma_{sym}(p\text{-xylene}^C,D_{2h}) = 0.14$ for the carbon structures, which reduces to $C_\sigma$ symmetry when the hydrogen atoms are included: $\sigma_{sym}(m\text{-xylene},C_{2v}) = 0.16$ and $\sigma_{sym}(m\text{-xylene},C_\sigma) = 0.06$, while $\sigma_{sym}(p\text{-xylene},D_{2h}) = 1.35$ and $\sigma_{sym}(p\text{-xylene},C_\sigma) = 0.09$. Similar results may be obtained with CSM, however, the current implementation of CSM only handles cyclic point groups and while these are correctly identified by the program[91], the reduction in symmetry (from e.g., $D_{nh}$ to $C_n$) that occurs when the protons are included are not shown.

Finally, 1,3,5-trimethylbenzene was investigated and the analysis is shown in Fig. 4g highlighting the effect of the proton positions on the symmetry. The in silico carbon structure has $D_{3h}$ symmetry and the molecular structure has $C_{3h}$ symmetry: $\sigma_{sym}(\mathbf{Q}^{UFF},C_{3h})$ = 5E-8. In contrast we find that the experimental structure has a significantly distorted symmetry as the carbon structure gives $\sigma_{sym}(\mathbf{Q}^C,D_{3h}) = 0.28$ and the experimental molecular structure gives $\sigma_{sym}(\mathbf{Q},C_{3h}) = 0.55$. Figure 4g shows the distortions that happen in the crystal structure to remove symmetry, which demonstrates the power inherent to the CSoM analysis when visualizing the outcome of the symmetry operations. As the current implementation of CSM can only handle cyclic symmetry point groups the symmetry reduction from $D_{3h}$ to $C_{3h}$ cannot be resolved using CSM[91].

### Polyaromatic hydrocarbons (PAHs)

Polyaromatic hydrocarbons are typically planar and highly symmetric molecules. We created idealized molecular structures using UFF in Avogadro using steepest descent and compared these to experimental crystal structures of the compounds. The CSoM analysis of six different polycyclic aromatics is provided in Table 1 and visualized in Fig. 4h for both optimized and crystal structures. In contrast to the methylbenzenes, the polyaromatic hydrocarbons only contain $sp^2$[7]-hybridized carbon atoms. Thus, these are more similar to benzene than toluene. That is up to a point, as the largest PAHs, like hexabenzocoronene, are no longer planar[132]. Table 1 shows the CSoM analysis of naphthalene, triphenylene, and hexabenzocoronene, illustrating this deviation from planarity.

The smaller PAHs all have $D_{nh}$ symmetry, as they are perfectly planar, the CSoM results are compiled in Table 1. The determined $\sigma_{sym}(\mathbf{Q},D_{nh})$ are low for the smaller PAHs. However, for triphenylene that has sterically congested hydrogens[132], a small twist can be observed in the experimentally determined structure[133]. As this twist is not found with the UFF method, the twist is readily observed in the overlay in Fig. 4h. The twist also manifests in a deviation from $D_{3h}$ with $\sigma_{sym}(\text{triphenylene}^C,D_{3h}) = 0.15$, while $C_{3v}$ without a mirror plane in the plane of the molecule has $\sigma_{sym}(\text{triphenylene}^C,C_{3v}) = 0.09$ and $\sigma_{sym}(\text{triphenylene}^C,C_2) = 0.06$. These values are for the carbon structure alone, if the molecular structure is considered we find $\sigma_{sym}(\text{triphenylene},D_{3h}) = 0.20$ and $\sigma_{sym}(\text{triphenylene},C_{3v}) = 0.13$ and conclude that triphenylene only has $C_2$ symmetry with $\sigma_{sym}(\text{triphenylene},C_2) = 0.07$. The same type of distortion leads hexabenzocoronene to have $D_{3d}$ symmetry instead of $D_{6h}$ symmetry.

### Transition metals, coordination polyhedra

With the demonstration of how the CSoM method can analyze organic molecules, we move to metal complexes. We start in the *d*-block with transition metal complexes, and we only use experimentally determined molecular structures. Here the SHAPE program has been used extensively to determine the structure of coordination polyhedra[103,112,113106], which we also refer to as the first or inner coordination sphere in a complex. Thus, we start with the CSoM analysis by investigating just the coordination polyhedron. Note that SHAPE does not differentiate between donor atoms in the coordination polyhedron. SHAPE only compares the shape or geometry formed by the coordinating donor atoms to selected reference polyhedra, thus we also used Polynator that finds the correct reference polyhedra, see

**Table 1 | Determined symmetry G and CSoM determined σsym-values of benzene, methylbenzene crystals, polyaromatic hydrocarbon with σsym-values of UFF optimized structure $Q^{UFF}$ in parenthesis**

| Q | Carbon only (CSoM) | Molecule (CSoM) | | Molecule (CSM) | ref |
|---|---|---|---|---|---|
| | G: $\sigma_{sym}$(Q,G) ($\sigma_{sym}$($Q^{UFF}$,G)) | G: $\sigma_{sym}$(Q,G) ($\sigma_{sym}$($Q^{UFF}$,G) | | G: S(Q,G) G: S($Q^{UFF}$,G) | |
| Benzene 1100158 | $D_{6h}$: 2.5E-3 (3.9E-8) | $D_{6h}$: 6.0E-3 (3.8E-8) | $C_2$: 2.0E-3 (2.0E-9) | $C_6$: 3.0E-3 (0.0) | 172 |
| Benzene 251255 | $D_{6h}$: 0.02 (3.9E-8) | $D_{6h}$: 0.06 (3.8E-8) | $C_2$: 0.03 (2.0E-9) | $C_6$: 0.03 (0.0) | 173 |
| Benzene 1100049 | $D_{6h}$: 0.23 (3.9E-8) | $D_{6h}$: 0.33 (3.8E-8) | $C_2$: 0.17 (2.0E-9) | $C_6$: 0.17 (0.0) | 174 |
| Toluene | $C_{2v}$: 0.016 (0.001) | $C_{2v}$: 1.41 (1.82) | $C_\sigma$: 0.053 (1E-6) | $C_2$: 1.74 (2.0) | 175 |
| *o*-xylene | $C_{2v}$: 0.03 (2E-9) | $C_{2v}$: 0.33 (2E-9) | $C_\sigma$: 0.04 (2E-9) | $C_2$: 0.14 (0.0) | 176 |
| *m*-xylene | $C_{2v}$: 0.08 (2E-6) | $C_{2v}$: 0.16 (8E-6) | $C_\sigma$: 0.06 (4E-6) | $C_2$: 0.03 (0.0) | 177 |
| *p*-xylene | $D_{2h}$: 0.14 (1E-3) | $D_{2h}$: 1.35 (1.71) | $C_{2h}$: 0.09 (2E-9) | $C_2$: 0.04 (0.0) | 178 |
| 1,3,5-trimethyl-benzene | $D_{3h}$: 0.28 (0.001) | $D_{3h}$: 1.77 (1.82) | $C_{3h}$: 0.55 (5E-8) | $C_3$: 0.18 (0.0) | 179 |
| Naphthalene | $D_{2h}$: 0.04 (2E-9) | $D_{2h}$: 0.04 (2E-9) | $C_{2h}$: 0.01 (2E-9) | n/a | 180 |
| Pyrene | $D_{2h}$: 0.02 (3E-9) | $D_{2h}$: 0.03 (3E-9) | $C_{2h}$: 0.01 (3E-9) | n/a | 181 |
| Triphenylene | $D_{3h}$: 0.15 (1E-6) | $D_{3h}$: 0.20 (1E-6) | $C_{3h}$: 0.13 (1E-7) | n/a | 133 |
| Perylene | $D_{2h}$: 0.01 (4E-4) | $D_{2h}$: 0.01 (1E-3) | $C_{2h}$: 6E-4 (1E-3) | n/a | 182 |
| Coronene | $D_{6h}$: 1E-3 (4E-8) | $D_{6h}$: 7E-4 (4E-8) | $D_{3d}$: 7E-4 (1E-10) | n/a | 178 |
| Hexabenzocoronene | $D_{6h}$: 4.8 (3.2) | $D_{6h}$: 5.4 (3.7) | $D_{3d}$: 0.02 (5E-3) | n/a | 132 |

Table S3. Either implementation of CShM has similar limitations, thus we limit our discussion to SHAPE.

The CSoM analysis accounts for both the position and the nature of the coordinating atoms. In this section we start by comparing the differences in the CShM and CSoM analyses, in the following section we show that it can be important to consider more than just the coordination atoms when determining the symmetry of a metal complex. As the symmetry depends on the coordination numbers, we differentiate based on these. The full analysis is compiled in Table 2, and the structures can be seen in Fig. 3 and the Supporting Information pages S4-S50.

Starting with CN = 4, SHAPE finds that $FeCl_4$ is a good match to the tetrahedron (P = T-4) with a CShM value of $\sigma_{ideal}(\mathbf{Q}^{inner}, \text{T-4}) = 0.23$. The CSoM analysis show that the symmetry is not $T_d$ as $\sigma_{sym}(\mathbf{Q}^{inner}, T_d) = 0.97$. Rather, iron(II) chloride is found to be $C_{2v}$ symmetric with $\sigma_{sym}(\mathbf{Q}^{inner}, C_{2v}) = 0.03$. Where SHAPE reports that iron(II) chloride has tetrahedral geometry—inferring $T_d$ symmetry, the CSoM analysis reports that iron(II) chloride has $C_{2v}$ symmetry. Which symmetry that determines the molecular properties is difficult to conclude without experimental testing, but if only SHAPE was used the symmetry may have been overestimated for iron(II) chloride.

The square planar nickel complex $[Ni(\mathbf{L})]^+$[134] also has CN = 4, but has three different types of coordinating nitrogen atoms that almost form the plane. This is confirmed by SHAPE as the best polyhedron is found to be the square (P = SP-4) with $\sigma_{ideal}(\mathbf{Q}^{inner}, \text{SP-4}) = 0.72$, which indicates $D_{4h}$ symmetry. The value found with SHAPE is small, but larger than the 0.32 found for iron(II) chloride. The symmetry of $Ni(\mathbf{L})]^+$ is not that of SP-4, as $\sigma_{sym}(\mathbf{Q}^{inner}, D_{4h}) = 1.4$. $C_{2v}$ symmetry is a better match with $\sigma_{sym}(\mathbf{Q}^{inner}, C_{2v}) = 0.56$, which is still large, and the structure may be considered asymmetric.

Figure 5a-c and Fig. 5f shows the CShM and CSoM analyses of the coordinating atoms of four complexes: $Fe^{II}Cl_4$ (CN = 4)[135], $[Ni(\mathbf{L})]^+$ (CN = 4)[134], $Fe^{III}Cl_6$ (CN = 6)[136], and $Fe(\mathbf{btz})_3^{3+}$(CN = 6)[137]. The two iron(III) complexes with CN = 6 are with SHAPE found to be closest to the octahedron (P = OC-6). SHAPE finds $\sigma_{ideal}(\mathbf{Q}^{inner}, \text{OC-6}) = 0.13$ for $Fe^{III}Cl_6$ which is significantly lower than the $\sigma_{ideal}(\mathbf{Q}^{inner}, \text{OC-6}) = 1.4$ determined for $Fe(btz)_3^{3+}$. The former complex may be considered more octahedral, and CSoM shows that this is reflected in the symmetry. $Fe^{III}Cl_6$ with $\sigma_{sym}(\mathbf{Q}^{inner}, O_h) = 0.08$ can be described with $O_h$ symmetry, while $Fe(\mathbf{btz})_3^{3+}$ with $\sigma_{sym}(\mathbf{Q}^{inner}, O_h) = 2.5$ cannot. However, the best CSoM result provides additional information. While $O_h$ can describe $Fe^{III}Cl_6$, the best CSoM result is for $D_{4h}$ symmetry with $\sigma_{sym}(\mathbf{Q}^{inner}, D_{4h}) = 0.01$. Whether this detail is important requires experimental investigations. $Fe(\mathbf{btz})_3^{3+}$, which did not have $O_h$ symmetry despite having an octahedral shape, can instead be shown to have $D_3$ symmetry with $\sigma_{sym}(\mathbf{Q}^{inner}, D_{3d}) = 0.01$. Note that no idealized polyhedron with 6 vertexes can be constructed with $D_{3d}$ symmetry, thus the correct geometry for the molecular structure for $Fe(\mathbf{btz})_3^{3+}$ cannot be determined with SHAPE and Polynator.

Detailed inspection of the CSoM output structures reveal that while the $Fe^{III}Cl_6$ does have slightly shorter distances to the axial donor atoms with $Fe\text{-}Cl_{ax} = 2.41$ Å and $Fe\text{-}Cl_{eq} = 2.55$ Å the complex can likely be regarded as $O_h$ symmetric. The perturbation to the structure is reflected in the better match to $D_{4h}$ symmetry.

A total of 25 structures were analyzed with both SHAPE, CSM, and the CSoM method. In many cases we found that the molecular structure had a different symmetry than the coordination polyhedron. Thus, we cannot investigate the relevant structures from the SHAPE results, which only consider the latter. With SHAPE, one general trend is worth noting: We observed that all structures with CN = 6 were found

**Table 2 | Determined symmetry G and CSoM determined σsym-values of metal complexes Q sorted by coordination number (CN). CShM values determined with SHAPE on the coordination polyhedral $Q^{inner}$, and σsym-values of the coordination polyhedra are included for comparison**

| Q | CN | Molecule | | | Coordination polyhedra | | | | ref. |
|---|---|---|---|---|---|---|---|---|---|
| | | Symmetry G | CSoM $\sigma_{sym}$(Q,G) | G: $\sigma_{sym}$(Q,G) | Symmetry G | CSoM σsym($Q^{inner}$,G) | CShM/SHAPE P: $\sigma_{ideal}$($Q^{inner}$,P) | G of P: $\sigma_{ideal}$($Q^{inner}$,G) | |
| Cu(**cyhdc**)(**MMA**)$_2$ | 4 | $C_i$ | 0.0 | $C_{2h}$: 2.6 | $C_{2h}$ | 0.0 | SP-4: 0.11 | $D_{4h}$: 34 | 183 |
| [Ni(**L**)]$^+$ | 4 | $C_\sigma$ | 5.1 | $C_{2v}$: 6.6 | $C_{2v}$ | 0.56 | SP-4: 0.72 | $D_{4h}$: 1.4 | 134 |
| $FeCl_4$ | 4 | $C_{2v}$ | 0.032 | - | $C_{2v}$ | 0.032 | T-4: 0.23 | $T_d$: 0.97 | 135 |
| $FeCl_6$ | **6** | $D_{4h}$ | 0.0 | - | $\boldsymbol{D_{4h}}$ | 0.0 | **OC-6**: 0.08 | $O_h$: 0.13 | 136 |
| Fe(**bpy**)$_3^{3+}$ | **6** | $C_2$ | 2.4 | - | $\boldsymbol{C_2}$$^a$ | 0.29 | **OC-6**: 0.91 | $O_h$: 15 | 184 |
| Fe(**phtmeimb**)$_2$ | **6** | $C_{2h}$ | 0.03 | $D_{3d}$: 0.83 | $\boldsymbol{D_{3d}}$ | 0.008 | **OC-6**: 0.094 | $O_h$: 0.17 | 185 |
| CuMn-MFU-4 (MOF) | **6** | $T_d$ | 0.0 | - | $\boldsymbol{T_d}$ | 0.0 | **OC-6**: 0.00 | $O_h$: 29 | 140 |
| Fe(**btz**)$_3^{3+}$ | **6** | $S_6$ | 0.01 | $D_3$: 0.064 | $\boldsymbol{D_3}$ | 0.064 | **OC-6**: 1.4 | $O_h$: 2.5 | 137 |
| $CdCl_2$15-**crown**-5 | 7 | $D_{5h}$ | 4.6 | - | $D_{5h}$ | 2.1 | PBPY-7: 1.5 | $D_{5h}$: 2.1 | 139 |
| W(**IMes**)(**H2BMes**)$_3$ | 8 | $C_2$ | 0.0 | - | $C_2$$^a$ | 0.0 | HH-9: 11 | $C_{2v}$: 1.5 | 186 |
| [Eu(**sst**)$_3$(THF)$_2$] | 5 | $C_3$ | 0.29 | $C_{3v}$: 0.61 | $C_{3v}$ | 0.15 | TBPY-5: 0.48 | $D_{3h}$: 0.31 | 152 |
| Dy(**bbpen**)X | 7 | $C_2$ | 0.0 | - | $C_2$$^a$ | 0.0 | PBPY-7: 2.05 | $D_{5h}$: 59 | 187 |
| Dy(**cy**$_3$PO)$_2$(H2O)$_5$ | 7 | $C_\sigma$ | 2.9 | - | $C_{5v}$ | 0.1 | PBPY-7: 0.142 | $D_{5h}$: 0.11 | 145 |
| Nd(**pc**)$_2$ | 8 | $C_4$ | 0.0 | $D_4$: 0.04 | $D_4$ | 0.0 | CU-8: 0.38 | $O_h$: 0.69 | 146 |
| Bu-**mim**[Eu(**tta**)$_4$] | 8 | $C_2$ | 2.2 | $D_2$: 58 | $D_2$ | 0.37 | SAPR-8: 0.36 | $D_{4d}$: 0.48 | 188 |
| Et-**mim**[Eu(**tta**)$_4$] | 8 | $D_2$ | 40 | $D_2$: 74 | $D_2$ | 0.26 | SAPR-8: 0.31 | $D_{4d}$: 0.57 | 188 |
| Tm(**DOTA**) | 8 | $C_4$ | 0.15 | - | $C_4$ | 0.02 | SAPR-8: 2.46 | $D_{4d}$: 0.69 | 162 |
| Dy(**DOTA**)($H_2O$) | 9 | $C_4$ | 0.08 | - | $C_4$ | 0.02 | CSAPR-9: 0.47 | $C_{4v}$: 0.49 | 162 |
| Eu(**oda**)$_3$ | 9 | $C_2$ | 0.13 | $D_3$: 1.10 | $D_3$ | 0.2 | TCTPR-9: 1.2 | $D_{3h}$: 1.34 | 189 |
| Nd(**oda**)$_3$ | 9 | $C_2$ | 0.72 | $D_3$: 1.30 | $D_3$ | 0.29 | TCTPR-9: 1.7 | $D_{3h}$: 8.8 | 85 |
| Nd($H_2O$)$_9$ Crystal | 9 | $C_3$ | 0.0 | -- | $C_3$$^a$ | 0.0 | TCTPR-9: 0.37 | $D_{3h}$: 0.19 | 147 |
| Nd($H_2O$)$_9$ DFT | 9 | $C_2$ | 3.13 | $C_{4v}$: 6.16 | $C_{4v}$ | 0.62 | TCTPR-9: 1.61 | $D_{3h}$: 2.3 | 138 |
| Dy(-**py**)$_2$ | 12 | $S_6$ | 0.01 | $D_{3d}$: 0.09 | $D_{3d}$ | 0.0 | IC12: 0.502 | $I_h$: 0.96 | 148 |
| $UP_5W_{30}O_{110}$ | 11 | $C_{5v}$ | 0.20 | - | $C_{5v}$ | 0.15 | JCPAPR-11: 1.157 | - | |
| β-Cm($PW_{11}O_{39}$) | 8 | $C_\sigma$ | 6.15 | $D_{4d}$: 11.0 | $D_{4d}$ | 0.63 | SAPR-8: .41 | $D_{4d}$: 0.63 | 110,190 |
| α-Cm($PW_{11}O_{39}$) | 8 | $C_2$ | 0.02 | $D_{4d}$: 6.34 | $C_2$$^a$ | 0.07 | SAPR-8 (0.333) | $D_{4d}$: 0.23 | 190 |
| $Np_2$(**mell**)($H_2O$)$_9$ (MOF) | 9 | $C_\sigma$ | 12.9 | - | $C_\sigma$$^a$ | 4.62 | MF-: 2.22 | $C_\sigma$: 4.62 | 191 |

$^a$The symmetry of this coordination polyhedra can also be determined using CSM, see Table S2 for details.

to be octahedral in the SHAPE analysis. This was not reflected in symmetry when the CSoM method was used, see the entries in bold in Table 2. While SHAPE did not differentiate these structures both Polynator and CSM did, see Tables S2 and S3, however, only CSoM determines the correct symmetry.

### Transition metal complexes

The electronic structure of a molecule or complex is determined by the position of the coordinating atoms alone. The orbitals on the coordinating atoms must be described[138]. This is readily achieved by analysing the symmetry of the molecular structure, rather than focusing exclusively on the coordinating polyhedron.

If we return to the possibly asymmetric [Ni(**L**)]$^+$ complex, where we analyzed the coordination polyhedra in Fig. 5a-c, the structure– shown in Fig. 3– with $\sigma_{sym}(\mathbf{Q}^{inner},C_{2v}) = 0.56$ could potentially have a $C_2$ axis or a mirror plane. When the molecular structure was investigated, we find $\sigma_{sym}(\mathbf{Q},C_{2v}) = 6.6$ and $\sigma_{sym}(\mathbf{Q},C_\sigma) = 5.1$, which show that the complex has no symmetry.

Figure 5d-f shows three complexes, for which we analysed the coordination polyhedral and the molecular structure. One CN = 7 cadmium(II) complex where 15-crown-5 provides 5 oxygen atoms as equatorial ligands and two chloride ions complete the coordination sphere on the axial positions[139]. The coordination polyhedron is best represented by the pentagonal bipyramid, and the symmetry is also found to best be described by $D_{5h}$. However, the $\sigma_{sym}$-values are large, and $CdCl_2$15-crown-5 must be considered to only have approximate symmetry at best. In contrast, CuMn-MFU-4, a metal organic framework with manganese as CN = 6 vertices, and where the manganese ions only have nitrogen atoms in the coordination polyhedron, is found to have perfect $O_h$ symmetry[140]. The coordination polyhedron has $\sigma_{sym}(\mathbf{Q}^{inner},O_h) = \sigma_{sym}(\mathbf{Q}^{inner},T_d) = 0.0$, however the molecular structure has $\sigma_{sym}(\mathbf{Q}^{inner},O_h) = 29$ and $\sigma_{sym}(\mathbf{Q}^{inner},T_d) = 0.0$. Thus, analysis of the coordination polyhedron provides a wrong description of the symmetry at the manganese atom, as molecular structure defines the molecule as $T_d$ symmetric. Note that SHAPE correctly defines the coordination polyhedron as an octahedron. Similarly, the [Fe(**btz**)$_3$]$^{3+}$ complex with CN = 6 is found to be an octahedron with SHAPE[137]. The CSoM method shows that the symmetry of the coordination polyhedron is in fact $D_{3d}$ with $\sigma_{sym}(\mathbf{Q}^{inner},D_{3d}) = 0.01$. However, if the molecular structure is analysed this is clearly wrong as we find $\sigma_{sym}(\mathbf{Q},D_{3d}) = 0.83$. The symmetry of [Fe(**btz**)$_3$]$^{3+}$ is actually $S_6$ with $\sigma_{sym}(\mathbf{Q},S_6) = 0.01$.

All the numbers, symmetry axis, and symmetry operated structures are automatically generated for all point groups when the CSoM method is used. Thus, we can readily determine the difference between the coordination polyhedron and the molecular structure of [Fe(**btz**)$_3$]$^{3+}$, see Fig. 5f. The difference between $D_{3d}$ and $S_6$ is six mirror planes, which are not present in the molecule. Note that $C_{2h}$ also describes the [Fe(**btz**)$_3$]$^{3+}$ well as a mirror plane exist orthogonal to the $C_2$-axis.

https://doi.org/10.1038/s41467-025-66054-7

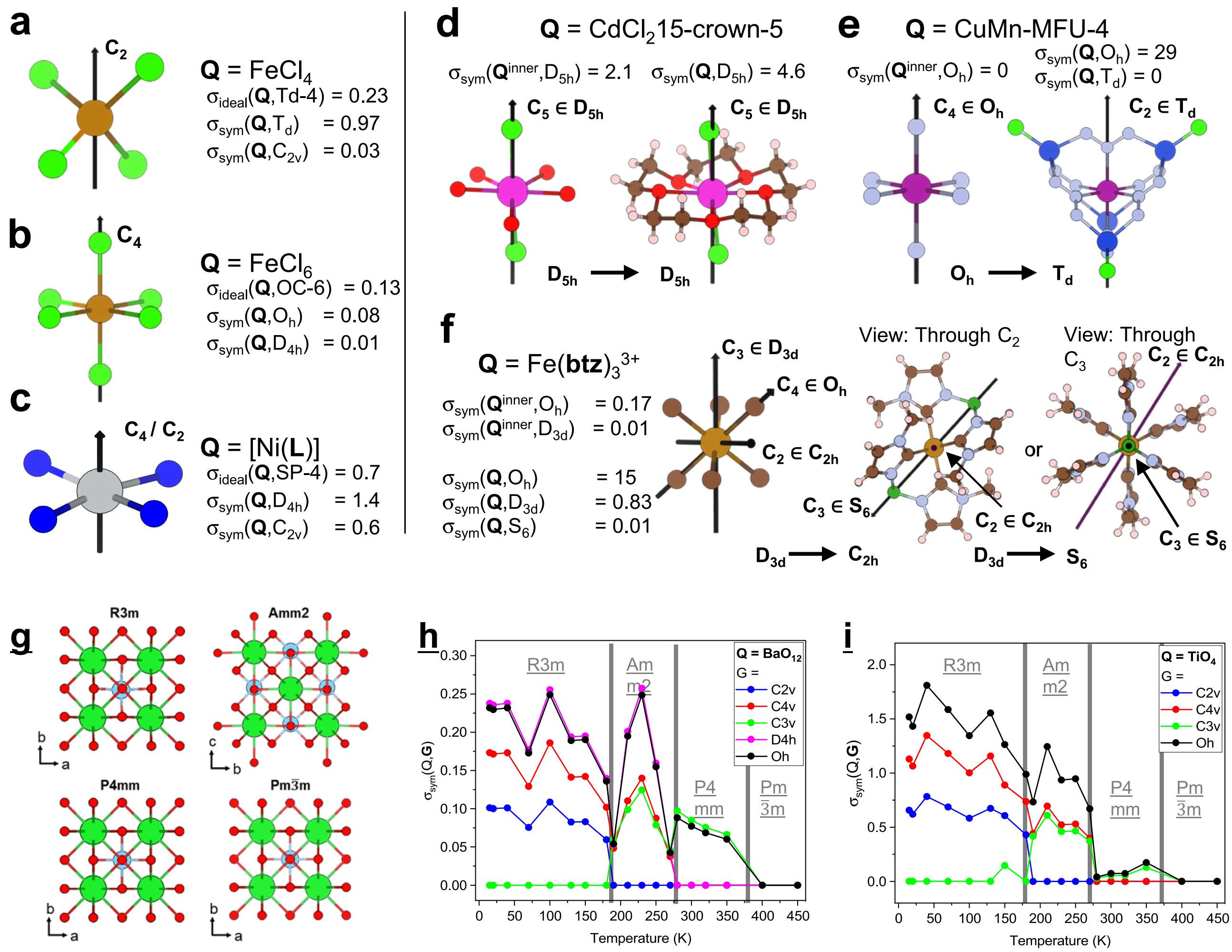


**Fig. 5 | Determination of the symmetry of the coordination polyhedra and molecules with CSoM for complexes and materials.** Comparing coordination polyhedra **P** determined with CShM using $\sigma_{ideal}$-values to the symmetry expressed with point groups **G** determined with CSoM via $\sigma_{sym}$-values for: **a** $Fe^{II}Cl_4$[135], **b** $[Ni(\mathbf{L})]^{+}$[134], **c** $Fe^{III}Cl_6$[136], **d** $CdCl_2$15-crown-5[139], **e** CuMn-MFU-4[140], **f** $Fe(\mathbf{btz})_3^{3+}$(CN = 6)[137]. **g** The four phases observed for the $BaTiO_3$ perovskite analyzed with CSoM for the of $BaO_{12}$ **h** and $TiO_4$ **i** centers in the $BaTiO_3$ perovskite as a function temperature. Structural data from reference[142] and [141]. Atom legend: carbon = brown, hydrogen = pink, oxygen = red, nitrogen = light blue, chlorine = green, titanium = ice blue, manganese = blue, iron = gold, copper = purple, nickel = light gray, cadmium = light purple, barium = bright green.

We explored a selection of transition metal complexes. The results are shown in Table 2. While the selected set of structures are not exhaustive, it is worth noting that none of the CSoM determined symmetries of the transition metal complexes are those determined by SHAPE for the coordination polyhedra.

### Phase transitions in $BaTiO_3$

The well-studied perovskite $BaTiO_3$ is known to have four different ferroelectric phases as a function of temperature. When changing the temperature of the perovskite from 10 to 450 K, three phase transitions have been observed between phases identified by the space groups R3m at low temperatures (<180 K), then Amm2 and P4mm, before forming Pm$\bar{3}$m symmetric lattice at high temperature (>400 K)[141,142]. The phase transitions have been reported through different structural parameters. These include the lattice constants, the Ti-O distances, the O-O distances, the equivalent isotropic parameters, the unit cell volumes, the electric polarization, the extinctions coefficient, and the strain broadening coefficients[142]. Recently, the structure of the coordination polyhedra was analyzed with an ellipsoidal analysis providing strong complementary value[118]. When considering the data we note that the Ti-O distances are not effected by temperature within each phase, but change abruptly between phases. Thus, we decided to determine the symmetry of the $BaO_{12}$ and $TiO_4$ centers in each phase, see Fig. 5g.

The CSoM analysis shows that symmetry of both metals can be evaluated through the temperature series, see Fig. 5h-i. In the R3m phase, the point group of both $BaO_{12}$ and $TiO_4$ is a perfect $C_{3v}$ symmetry, which turns into a perfect $C_{2v}$ in the Amm2 phase. At higher temperatures the symmetry is perfect $D_{4h}$ for $BaO_{12}$ and perfect $C_{4v}$ for $TiO_4$ in the P4mm space group, and both adopt perfect $O_h$ symmetry in the Pm$\bar{3}$m phase. In addition to determining the symmetry, the analysis show that the $\sigma_{sym}$-values change continuously as a function of temperature. This indicates that the overall symmetry in the perovskite increases with temperature. An insight which was not obvious before using CSoM[118,141,142].

### Lanthanide complexes

We started exploring molecular structures in the *f*-block, with the perception that our understanding for coordination numbers 4 to 6/7 was well established. A statement that holds for coordination polyhedra, but not for molecular structures, see above. Lanthanide complexes are typically of coordination numbers from 7 to 10, and finding the correct reference polyhedra is difficult. As we will show below using SHAPE in the *f*-block gives questionable results at best[111,143,144]. In contrast, we show that CSoM is both fully automated and robust.

Figure 6a-d shows the analysis of four lanthanide complexes with CN = 7, 8, 9, and 12. The Dy(**bbpen**)X complex with CN = 7 is best described by PBPY-7 using SHAPE with $\sigma_{ideal}(\mathbf{Q}^{inner},\text{PBPY-7}) = 2.05$, see Fig. 6a[145]. However, with $\sigma_{sym}(\mathbf{Q}^{inner},D_{5h}) = 69$ the symmetry of the coordination polyhedron is not $D_{5h}$ symmetric as PBPY-7 would suggest, but $C_2$, which is also the symmetry of the molecule: $\sigma_{sym}(\mathbf{Q},C_2) = 0.0$. The Nd(**pc**)$_2$ complex has CN = 8 and has a coordination polyhedron that is an almost perfect match to the cube polyhedron $\sigma_{ideal}(\mathbf{Q}^{inner},\text{CU-8}) = 0.38$, see Fig. 6b[146]. The $\sigma_{sym}$-values reveal a significant distortion from $O_h$ symmetry with $\sigma_{sym}(\mathbf{Q}^{inner},O_h) = 0.7$. Instead, the coordination polyhedron of Nd(**pc**)$_2$ is found to have $D_4$ symmetry with $\sigma_{sym}(\mathbf{Q}^{inner},D_4) = 0.0$. This symmetry is enforced by the molecular structure that has $\sigma_{sym}(\mathbf{Q},O_h) = 8.0$ and $\sigma_{sym}(\mathbf{Q},D_4) = 0.04$. Nd(**oda**)$_3$[147], with CN = 9, is matched to the TTP polyhedron by SHAPE $\sigma_{ideal}(\mathbf{Q}^{inner},\text{TCTPR-9}) = 1.7$, see Fig. 6c. However, the symmetry is not $D_{3h}$ with $\sigma_{sym}(\mathbf{Q}^{inner},D_{3h}) = 8.8$, but closer to $D_3$ with $\sigma_{sym}(\mathbf{Q}^{inner},D_3) = 0.3$. The $\sigma_{sym}$-value is high and scrutiny of the CSoM analysis shows that the structure is found to be better described by $C_2$ symmetry than $C_3$ symmetry. Both are subgroups to $D_3$. With a minimum $\sigma_{sym}$-value value of 1.3, the molecular structure of the complex is considered to be significantly distorted from $D_3$ symmetry. The last complex is Dy(**tp-py**)$_2$[148], a complex with CN = 12, see Fig. 6d. The coordinating atoms form an icosahedron with $\sigma_{ideal}(\mathbf{Q}^{inner},\text{IC-12}) = 0.5$, with an acceptable match to a distorted $I_h$ symmetry $\sigma_{sym}(\mathbf{Q}^{inner},I_h) = 0.96$. This value is a little large, thus the coordination polyhedral should be described as having $D_{3d}$ symmetry with $\sigma_{sym}(\mathbf{Q}^{inner},D_{3d}) = 0.001$ instead. The molecular structure is equally well described by $S_6$ and $D_{3d}$ symmetry with $\sigma_{sym}(\mathbf{Q},D_{3d}) = 0.10$ and $\sigma_{sym}(\mathbf{Q},D_{3d}) = 0.09$. As $S_6$ does not have any mirror planes perpendicular to the principal axis the transition probabilities in e.g., optical spectroscopy will be very different depending on whether the molecular structure is best described as $S_6$ or $D_{3d}$.

The analysis of the lanthanide complexes, emphasize the conclusion above: While the SHAPE can be a measure to identify a coordination polyhedron, these very often provide a poor description of the symmetry and molecular structure of the complex. In Table 2 SHAPE identify the correct structure in 1/25 cases, and in 11/25 cases the coordination polyhedron and molecular structure is of the same symmetry.

Thus far, we have only touched on how the CSoM analysis can be used to visually compare symmetries and structures, and how the analysis can be used to determine symmetry and distortions from symmetry. The CSoM method is particularly strong when used along experimental data as in the examples below.

## A series of tris(trimethylsilyl)siloxidelanthanide(III) complexes

The series of lanthanide ions has been explored by preparing isostructural complexes, looking for structural changes induced by the lanthanide contraction[147,149–151]. Along this vein, a series of CN = 5 Ln(**sst**)$_3$THF$_2$ complexes was reported by Boyle and coworkers[152]. In the report the structure of the complexes were described as having tricapped bipyramid (TBPY-5) coordination polyhedra, which indicates $D_{3h}$ symmetry. For the lighter, larger lanthanide(III) ions the description with TBPY-5 is worse than it is for the heavier, smaller lanthanide(III) ions. We revisited the data using the CSoM method. The results are shown in Fig. 6f, and the data for the Eu(III) complex is shown in Table 2.

As $0.3 < \sigma_{ideal}(\mathbf{Q}^{inner},\text{TBPY-5}) < 0.6$ SHAPE reports that the complexes can be described with the TBPY-5. Similarly, the CSoM method reports that with $\sigma_{sym}(\mathbf{Q}^{inner},C_3) \approx 0.2$, all the complexes have coordination polyhedra that are only slightly distorted from $C_3$ symmetry.

Considering the molecular structure Fig. 6f shows how the $\sigma_{sym}(\mathbf{Q},C_3)$-value clearly show how the difference in size from La(III) to Lu(III) impacts the distortion of symmetry of the complexes. This example illustrates how CSoM can improve the analysis of isostructural complexes based on the molecular structure[153,154].

## Ln.DOTA

Lanthanide(III) complexes of the **DOTA** ligand (**DOTA** = 1,4,7,11-tetraazacyclododecane-1,4,7,11-tetraacetic acid) are by far the most studied lanthanide(III) complexes[155–161]. Ln.**DOTA** crystallize with and without a capping water molecule resulting in complexes with CN = 8 and CN = 9. Figure 6g shows a CN = 8 Tm.**DOTA** and a CN = 9 Dy.**DOTA**($H_2O$) complex i.e., **DOTA** complexes of two trivalent lanthanides that crystallized with and without a capping water molecule[162]. As the coordination number is different, the CShM method provides different results describing the Tm.**DOTA** complex as the square antiprismatic polyhedron SAPR-8; $\sigma_{ideal}(\mathbf{Q}^{inner},\text{SAPT-8}) = 2.46$, and the Dy.**DOTA**($H_2O$) complex as the capped square antiprismatic polyhedron CSAPR-9; $\sigma_{ideal}(\mathbf{Q}^{inner},\text{CSAPR-9}) = 0.47$. Thus, the two coordination polyhedra cannot be compared using SHAPE. However, the CSoM method shows that both complexes have $C_4$ symmetry, both with $\sigma_{sym}(\mathbf{Q}^{inner},C_4) = 0.02$.

In this particular example, the electronic properties of Dy.**DOTA**($H_2O$) were studied as a function of the molecular structure, thus only the CSoM method is relevant. Considering the molecular structure, Tm.**DOTA** almost retains $C_4$ symmetry with $\sigma_{sym}(\mathbf{Q},C_4) = 0.15$. The molecular structure of the Dy.**DOTA**($H_2O$) complex has $C_4$ symmetry if the protons on the capping water molecule are ignored: $\sigma_{sym}(\mathbf{Q},C_4) = 0.08$. If the protons on the capping water molecule are included, the symmetry is broken: $\sigma_{sym}(\mathbf{Q}',C_4) = 0.85$. The importance of the orientation of the water molecule was recognized by Le Guennic and coworkers[163,164], who explored the effect on the electronic structure in silico. They found that the electronic ground state and the first excited state cross twice within a 180° rotation of the capping water molecule[163]. Which translated into significant changes in the magnetic properties of the complex. Determining the symmetry of the Dy.**DOTA**($H_2O$) complex with CSoM as the capping water is rotated, we see a periodic variation in the $\sigma_{sym}(\mathbf{Q},C_{4v})$, $\sigma_{sym}(\mathbf{Q},C_4)$, $\sigma_{sym}(\mathbf{Q},C_{2v})$, and $\sigma_{sym}(\mathbf{Q},C_2)$ values, see Fig. 6g. Interestingly, each of the four $\sigma_{sym}$-values are found to behave differently: $\sigma_{sym}(\mathbf{Q},C_4)$ is found to be invariant to the rotation, and $\sigma_{sym}(\mathbf{Q},C_2)$ varies with the full rotation of the water molecule with a period of 360°. The $\sigma_{sym}(\mathbf{Q},C_{2v})$-value has a period of 90°, while for the $\sigma_{sym}(\mathbf{Q},C_{4v})$-value it is 45°. The electronic properties determined by Le Guennic and coworkers showed a period of 90°, suggesting that the electronic properties of the Dy.**DOTA**($H_2O$) complex are governed by $C_{2v}$ symmetry[163,164].

## Europium luminescence

The lines observed in the emission spectrum of Eu(III) are routinely correlated to the perceived symmetry of a complex based on symmetry arguments[82,165,166]. The emissive term $^5D_0$ contains a single electronic state, while the final terms $^7F_0$, $^7F_1$, and $^7F_2$ contain 1, 3, 5 electronic states. The arguments for the three bands $^5D_0 \rightarrow {}^7F_J$ (J = 0, 1, and 2) can be stated briefly as:

- Without symmetry, 1, 3, and 5 lines (the maximum) will be observed in the three bands.
- In cubic symmetry groups (e.g., $O_h$ and $T_d$) 1 and 2 lines will be observed in $^5D_0 \rightarrow {}^7F_1$ and $^5D_0 \rightarrow {}^7F_2$ bands, respectively.
- In hexagonal or trigonal symmetry (e.g., $D_{3h}$ and $C_3$) 2 and 3 lines will be observed in $^5D_0 \rightarrow {}^7F_1$ and $^5D_0 \rightarrow {}^7F_2$ bands, respectively.
- In tetragonal symmetry (e.g., $C_4$ and $D_{4d}$) 2 and 4 lines will be observed in $^5D_0 \rightarrow {}^7F_1$ and $^5D_0 \rightarrow {}^7F_2$ bands, respectively.
- If the point group symmetry has a mirror plane or a $C_2$ axis perpendicular to the main symmetry axis the $^5D_0 \rightarrow {}^7F_0$ line will disappear. In cyclic groups (e.g., $C_n$) a single line will be observed.

 https://doi.org/10.1038/s41467-025-66054-7

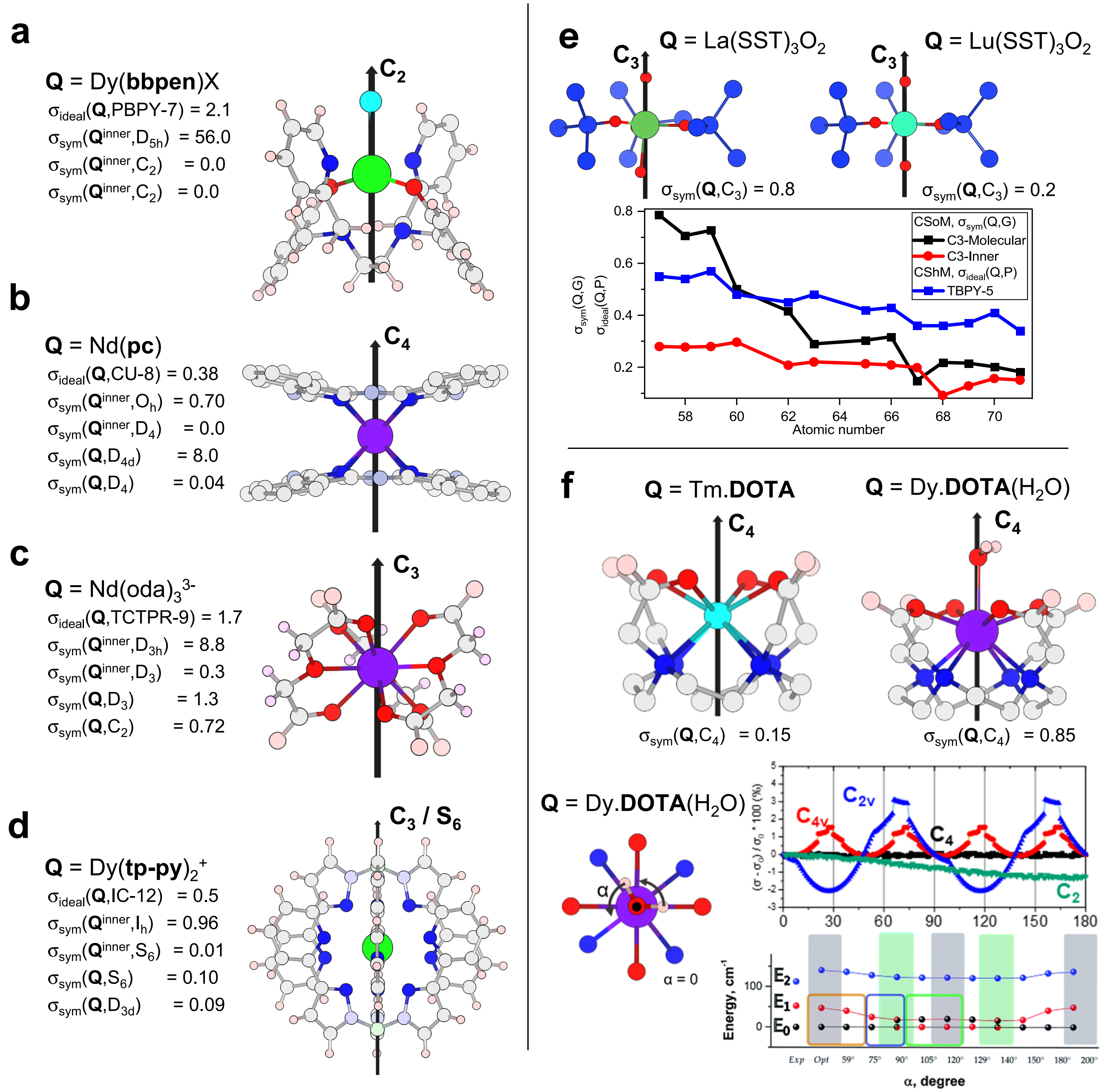


**Fig. 6 | Lanthanide complexes analysed with SHAPE v2.1 (CShM) and CSoM providing σideal-values for a selected polyhedra P and $\sigma_{sym}$-values for selected point groups G. a** Dy(**bbpen**)X[145], **b** Nd(**pc**)$_2$[146], **c** Nd(**oda**)$_3$[147], **d** Dy(**tp-py**)$_2$[148]. **e** CSoM investigation of the structure of Ln(**sst**)$_3$THF$_2$ series. **f** CSoM investigation of the structure of Tm.**DOTA** and Dy.**DOTA**($H_2O$) and the electronic structure of Dy.**DOTA**($H_2O$) as a function of symmetry[163]. Atom legend: carbon ligand backbone = pale gray, silicon ligand backbone = blue, hydrogen = pink, oxygen = red, aromatic nitrogen = light blue, aliphatic nitrogen = dark blue, lanthanum = green, neodymium = deep purple, dysprosium = bright green, thulium= aquamarine, lutetium = pale green. Images in panel **f** used with permission of Royal Society of Chemistry, from Chemical Science, M. Briganti *et al*, 10, 7233-7245, 2019; permission conveyed through Copyright Clearance Center, Inc.

Figure 7 displays the emission spectra of the bands $^5D_0\rightarrow{}^7F_J$ (J = 0, 1, and 2) for ten Eu(III) complexes that we selected to analyze using CShM and CSoM. A summary of the analysis is compiled in Table S4.

In Fig. 7a a crystalline phosphor is shown, in this material Eu(III) has cubic symmetric with $\sigma_{sym}$(Eu:Ba$_2$MgWO$_6$,$T_d$) = 0.0 aligning with the observation that a single line is observed for $^5D_0\rightarrow{}^7F_1$ while no line is observed for $^5D_0\rightarrow{}^7F_0$. In Fig. 7b the symmetry of Eu(III) in solid LnOCl is $C_{4v}$ and as expected a very intense $^5D_0\rightarrow{}^7F_0$ line and two lines within $^5D_0\rightarrow{}^7F_1$ were observed.

When moving to molecular materials in Figs. 7c and 7d the results show that the symmetry of the molecular structure is important. While first coordination sphere is described by $C_4$ symmetry with $\sigma_{sym}$(Eu.**DOTA**($H_2O$),$C_4$) = 0.1 and $\sigma_{sym}$(Eu.**L'**($H_2O$),$C_4$) = 0.3, thevalues for the molecular structures $\sigma_{sym}$(Eu.**L'**($H_2O$),$C_4$) = 43.7 and $\sigma_{sym}$(Eu.**DOTA**($H_2O$),$C_4$) = 0.5 show that only Eu.**DOTA**($H_2O$) has $C_4$ symmetry. The spectra confirm this finding as the splitting of the $^5D_0\rightarrow{}^7F_1$ band increase from 2 lines in Eu.**DOTA**($H_2O$) split into 3 three lines in Eu.**L'**($H_2O$), which corresponds to the reduction in symmetry from $C_4$ to $C_2$. Figure 7d shows that the maximum number of lines can be observed in Eu$_2$(SO$_4$)$_6$ that has perfect $C_2$ symmetry.

Considering trigonal groups, Eu(**dpa**)$_3$ has a molecular structure with perfect $D_3$ symmetry and the corresponding spectra in Fig. 7f show no $^5D_0\rightarrow{}^7F_0$ line and two lines in $^5D_0\rightarrow{}^7F_1$. Figure 7g shows the

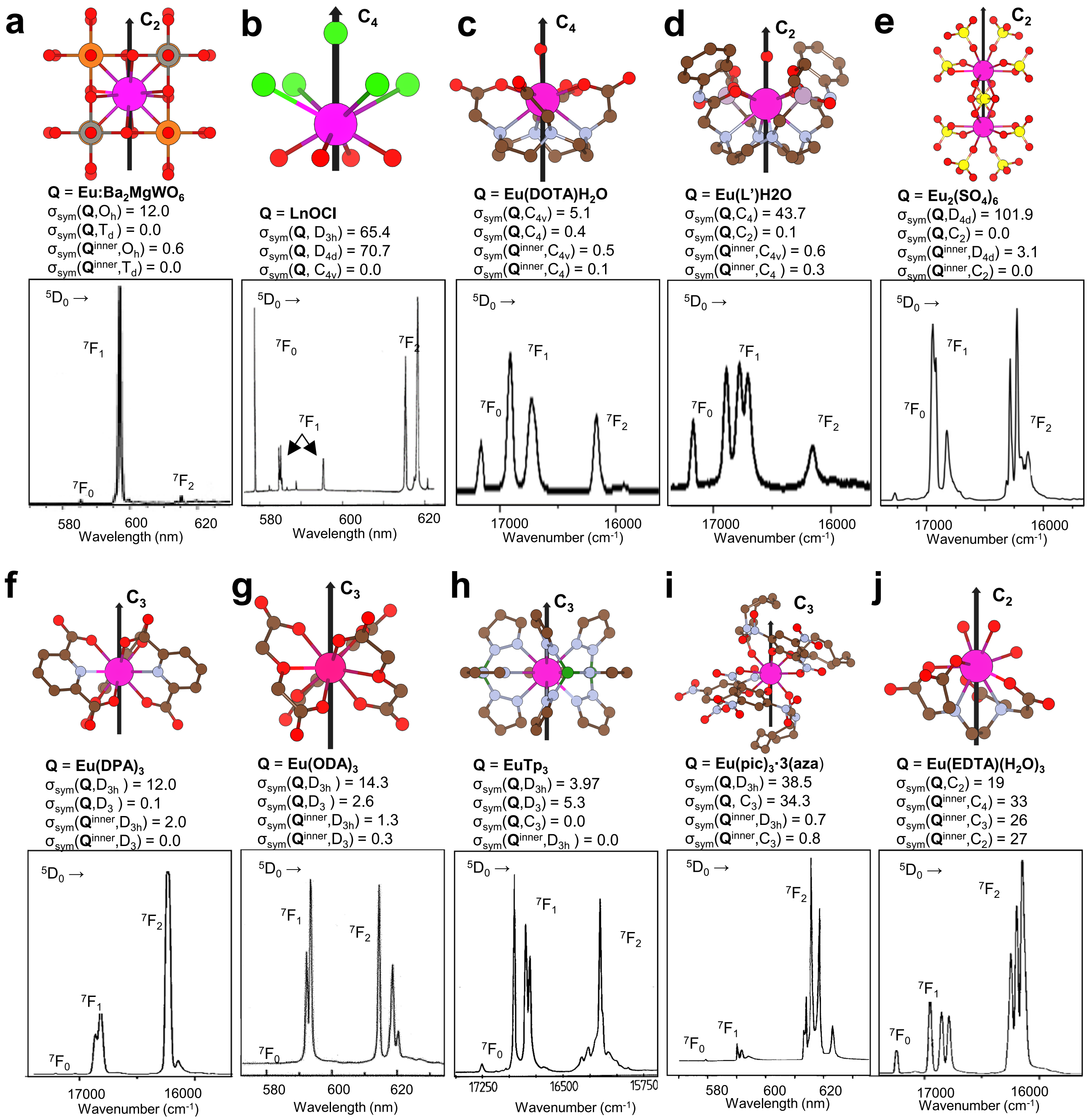


**Fig. 7 | Emission spectra and a CSoM analysis of relevant point groups for 10 Eu(III) complexes, σsym-values for selected point groups G are reported for the molecular structure Q and for the coordinating atoms $Q^{inner}$.** The structures and spectra are redrawn from literature data: **a.** Eu:$Ba_2MgWO_6$[196], **b.** $Eu_2(SO_4)_6$[111], **c.** EuOCl[197], **d.**Eu(**pic**)$_3$·3(aza)[198], **e.** Eu.**DOTA**($H_2O$)[199], **f.** Eu.**L'**($H_2O$)[199], **g.** Eu(**oda**)$_3$[200], **h.** Eu(**dpa**)$_3$[201], **i.** $EuTp_3$[202], and **j.** Eu.**edta**($H_2O$)$_3$[201]. Atom legend: europium = purple, carbon = brown, oxygen = red, nitrogen = light blue, chlorine = green, sulfur = yellow, barium = orange, tungsten = dark orange. Hydrogen atoms are omitted for clarity.

spectra Eu(**oda**)$_3$ with near perfect $D_3$ symmetry of the coordination atoms and a distorted molecular structure. While the spectra of Eu(**dpa**)$_3$ and Eu(**oda**)$_3$ are close to identical, the Eu(**oda**)$_3$ corresponding spectra appear to be split slightly more due to the lack of symmetry in the molecular structure. It should be noted that the CShM value for both TTP and cSAP are identical and small for both Eu(**oda**)$_3$ and Eu(**dpa**)$_3$ and CSM cannot identity the $D_3$ symmetry.

Figure 7h shows the Eu(**tp**)$_3$, with an inner coordination sphere with $D_{3h}$ symmetry, and a molecular structure with $C_3$ symmetry. This is expressed in the spectra as the $^5D_0 \rightarrow {}^7F_0$ line appear. The importance of considering the molecular structure is enforced when considering Eu(**pic**)$_3$·3**(aza)** in Fig. 7i where the CSoM values for the coordinating atoms are $\sigma_{sym}(\mathbf{Q}^{inner}, D_{3h})$= 0.7, but the maximum amount of lines that are observed in the spectra shows that the europium center does not have symmetry, supported by the lack of symmetry in the molecular structure. The last example is the completely asymmetric Eu.**edta**($H_2O$)$_3$ complex shown in Fig. 7j, where the $^5D_0 \rightarrow {}^7$F0 line is clearly observed next to three well-separated lines in the $^5D_0 \rightarrow {}^7F_1$ band.

## Nd(III) electronic structure

The crystal field splitting of Nd(III) are not as readily explained as the crystal field splitting of Eu(III)[84]. However, the electronic structure of

https://doi.org/10.1038/s41467-025-66054-7

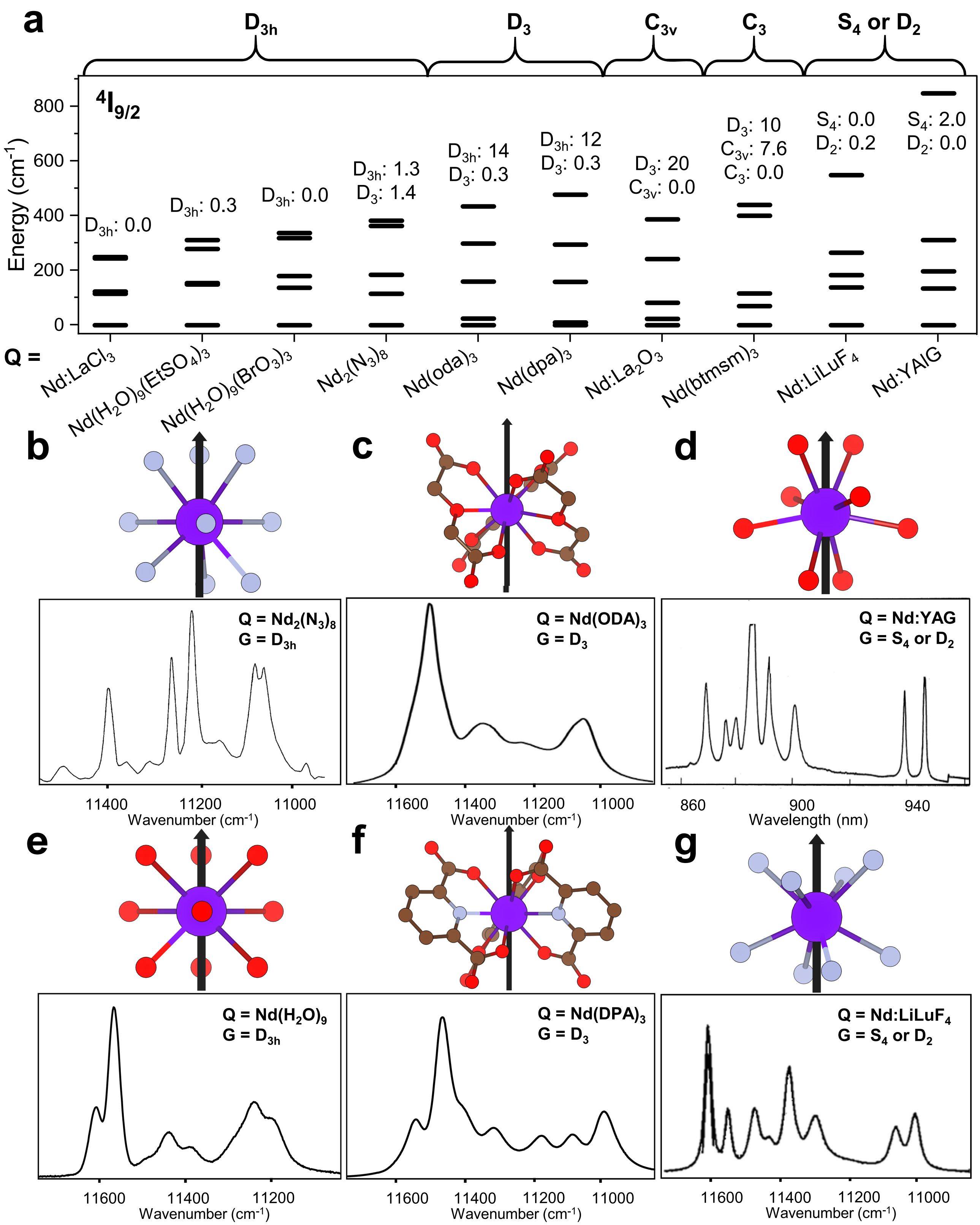


**Fig. 8 | The effect of symmetry on the Nd(III) electronic structure and emission spectra.** **a** Crystal field splitting in the lowest energy term $^4I_{9/2}$ of Nd(III) in 10 different environments: Nd:$LaCl_3$[203], $Nd(H_2O)_9(EtSO_4)_3$[204], $Nd(H_2O)_9(BrO_3)_3$[205], $Nd_2(N_3)_8$[206], Nd(**oda**)$_3$[85], Nd(**dpa**)3 (Note: CSoM values have been taken from the Eu(**dpa**)$_3$ structure as the crystal field splitting has only been determined in solution)[167], Nd:$La_2O_3$[207], Nd(**btmsm**)$_3$[208], Nd:$LiLuF_4$[168], and Nd:YAG[169]. The energy levels were determined with emission spectroscopy for each molecular structure **Q** and the point group **G** was determined with a CSoM analysis. The relevant $\sigma_{sym}$(**Q**,**G**)-values are shown on the plot. The emission spectra of the $^4F_{3/2} \rightarrow {}^4I_{9/2}$ band for selected molecular structures **Q** with their CSoM determined point group **G**, are shown for **b** $Nd_2(N_3)_8$, **c** Nd(**oda**)$_3$, **d** Nd:YAG, **e** $Nd(H_2O)_9$, **f** Nd(**dpa**)$_3$, and **g** Nd:$LiLuF_4$ complexes, the spectra are redrawn from earlier work[85,137,167–169,206]. Atom legend: neodymium = deep purple, carbon = brown, oxygen = red, nitrogen = light blue. Hydrogen atoms are omitted for clarity.

the central ion must still be influenced by the symmetry of the environment. Figure 8a shows the crystal field splitting of the $^4I_{9/2}$ term of Nd(III) in ten different materials, and trends emerge when the crystal field splitting diagrams are sorted by symmetry. All $D_{3h}$ symmetric Nd(III) ions share a crystal field splitting where the levels are grouped in a (1, 2, 2) distribution. Reducing the symmetry to $D_3$, $C_{3v}$, or $C_3$ the crystal field splitting is different, which is evident in the Nd(III) emission spectrum. Considering Fig. 8a, we note that the two molecular structures with $D_3$ symmetry, Nd(**oda**)$_3$[85] and Nd(**dpa**)$_3$[167], share the same distribution of (2, 1, 1, 1). In the $S_4$ symmetric Nd:LiLuF$_4$[168], and the $D_2$ symmetric Nd:YAG[169] materials we find a similar distribution of levels. The CSoM analysis determined that the Nd:LiLuF$_4$[168] structure has slightly distorted $D_2$ symmetric with $\sigma_{sym}(\mathbf{Q}^{inner},D_2) = 0.2$ and that the Nd:YAlG has distorted $S_4$ symemtry with $\sigma_{sym}(\mathbf{Q}^{inner},S_4) = 2.0$, and must consider that these distortions might be of a magnitude where the effective symmetry of both structures are the same. In Fig. 8b-g the emission spectra showing the Nd(III) $^4F_{3/2} \rightarrow {}^4I_{9/2}$ band are shown for six materials. This data is condensed in the crystal field splitting diagram, but cursory inspection of the band shapes highlights that minute differences and similarities in symmetry are readily translated to spectral features in lanthanide(III) luminescence.

## Discussion

Any quantitative method needs a yardstick and when it comes to symmetry a cut-off must indicate the presence and absence of symmetry. A real cut-off must be based on experimentally determined structures and symmetry dictated properties e.g., splitting of signals in NMR, EPR etc. The analysis here suggests the following cut-off values based on input data. The symmetry analysis of water suggests a numerical cut-off of $\sigma_{sym} = 1E\text{-}4$ for symmetry when using CSoM on in silico data of small molecules. For larger molecular structures a cut-off would appear to be $\sigma_{sym} < 0.1$.

Considering the CSoM analysis for *ortho*-xylene we determine $\sigma_{sym}(o\text{-xylene}^C,C_{2v}) = 0.03$ for the carbon structure, while we find $\sigma_{sym}(o\text{-xylene},C_{2v}) = 0.33$ and $\sigma_{sym}(o\text{-xylene},C_\sigma) = 0.04$ for the molecular structure. Using the cutoff of $\sigma_{sym} < 0.1$ the symmetry of *ortho*-xylene is assigned as $C_\sigma$ for the molecule and $C_{2v}$ for the carbon structure. Similarly for *meta*- and *para*-xylene: with $\sigma_{sym}(m\text{-xylene},C_{2v}) = 0.16$ and $\sigma_{sym}(m\text{-xylene},C_\sigma) = 0.06$ the symmetry of *meta*-xylene was determined to be $C_\sigma$, and with while $\sigma_{sym}(p\text{-xylene},D_{2h}) = 1.35$ and $\sigma_{sym}(p\text{-xylene},C_\sigma) = 0.09$ the symmetry of *para*-xylene was determined to be $C_\sigma$. These numbers are not the solution, but the beginning of a yardstick.

In summary, we find that there is nothing more important in chemistry than molecular structure. However, we do not have a measure that allows us to talk about molecular structure at the level of detail that experimental and theoretical methods now provide. Here, we have shown the limits of current methods for determining and discussing molecular structure and symmetry, and we have highlighted the power of the Continuous Symmetry operation Measure (CSoM) methodology through a series of case studies spanning the symmetry in the molecular structure of water, polycyclic aromatic hydrocarbons, classical transition metal complexes, single molecule magnets, and lanthanide complexes. We conclude that the CSoM method works across all inputs, automatically ensures that the correct molecular coordinate system is used, and provides both numerical and visual outputs.

To demonstrate applicability of CSoM and demonstrate the relevance of symmetry as a relative measure we used lanthanide(III) luminescence. The symmetry determined with the CSoM method was used to rationalize the crystal field splitting and band shapes of Nd(III) and Eu(III) emission spectra, and we noted that the CSoM results reflect both the physical distortion from symmetry and the resulting molecular properties.

Using CSoM values requires a yardstick and based on the results presented here we conclude that for in silico molecules symmetry is present if $\sigma_{sym} \leq 1E-4$. While experimentally determined molecular structures often carry a larger error and are symmetric if $\sigma_{sym} < 0.1$. In general, $1 \leq \sigma_{sym} < 5$ reflects structures that are distorted or heavily distorted from a symmetry. As the size of the molecular structure is relevant, in particular for properties arising from single atoms e.g., transition metal or lanthanide centers, the $\sigma_{sym}$-values for larger structures may have less strict symmetry cut-off values. These values are to be considered as a starting point only, as they are derived from a small set of structures without considering a specific experimentally observable property. In order to provide physical meaningful CSoM cut-off values for symmetry, a specific symmetry dependent property must be declared, and the proposed values should be tested on a significant sample set. Here, we verified the selection rules in europium(III) luminescence and begun to analyse the electronic structure of neodymium(III).

In conclusion, we propose a path to making comparative studies across chemistry and material science easier. First, we propose that chemists use the IUPAC definition of molecular entities when correlating molecular structure to molecular properties. Second, we propose the CSoM methodology as a unifying method to quantify and analyze molecular structure. And third, we propose that we start using σsym-values in combination with experimental results to set a σsym cut-off for when symmetry dictates an observed property.

## Methods

### Structure retrieval

All structures have been manually retrieved as '.cif' files from the ccdc database. The '.cif' files have been used to construct the '.xyz' files manually in the software VESTA v3. The distinction between molecular structure and first coordination sphere has been done manually.

### In silico structures

In silico structures have been geometry optimized from the relevant crystal structures retrieved, see above, in Avogadro[170] using UFF force fields[171]. For the geometry optimizations 500 steps have been used with the steepest descent and a convergence threshold of 10E-10.

### CSM

The Continuous Symmetry Measure (CSM) program has been used on the webpage csm.ouproj.org.il. The implementation is reported in ref. 91.

### CShM

The Continuous Shape Measure (CShM) has been used with the SHAPE v2.1 program. The implementation is reported in ref. 103.

### CSoM

The Continuous Symmetry operation Measure (CSoM) has been used on the structure files in the '.xyz' data format. The program was downloaded from (https://github.com/VRMNielsen/Continous-Symmetry-Operation-Measure-Program). The implementation is reported in ref. 2.

### Polynator

The Polynator program has been used as a different shape measure. The program has been downloaded from (https://journals.iucr.org/j/issues/2023/06/00/jl5072/index.html). The implementation is reported in ref. 117.

## Data availability

The structures used and data generated in this study have been deposited in the FigShare database (https://doi.org/10.6084/m9.figshare.29382491.v1). Visualizations of the data are included in the supporting information page S6-S50. All data are available from the corresponding author upon request.

## Code availability

The CSoM code used in this article have been uploaded to GitHub (https://github.com/VRMNielsen/Continous-Symmetry-Operation-Measure-Program). Polynator (https://journals.iucr.org/j/issues/2023/06/00/jl5072/index.html).

## References


1. Sheka, E. F., Razbirin, B. S. & Nelson, D. K. Continuous symmetry of C60 fullerene and its derivatives. *J. Phys. Chem. A* **115**, 3480–3490 (2011).
2. Nielsen, V. R. M., Le Guennic, B. & Sørensen, T. J. Evaluation of point group symmetry in lanthanide(III) complexes: a new implementation of a continuous symmetry operation measure with autonomous assignment of the principal axis. *J. Phys. Chem. A* **128**, 5740–5751 (2024).
3. Pilati, T. & Forni, A. SYMMOL: a program to find the maximum symmetry group in an atom cluster, given a prefixed tolerance. *J. Appl. Crystallogr.* **31**, 503–504 (1998).
4. Knowles, P. J. The determination of point groups from imprecise molecular geometries. *J. Math. Chem.* **60**, 161–171 (2022).
5. Alon, G. & Tuvi-Arad, I. Improved algorithms for symmetry analysis: structure preserving permutations. *J. Math. Chem.* **56**, 193–212 (2018).
6. Dryzun, C. & Avnir, D. Generalization of the continuous symmetry measure: the symmetry of vectors, matrices, operators and functions. *Phys. Chem. Chem. Phys.* **11**, 9653–9666 (2009).
7. Pinsky, M., Zait, A., Bonjack, M. & Avnir, D. Continuous symmetry analyses: C(nv) and D(n) measures of molecules, complexes, and proteins. *J. Comput Chem.* **34**, 2–9 (2013).
8. Largent, R. J., Polik, W. F. & Schmidt, J. R. Symmetrizer: algorithmic determination of point groups in nearly symmetric molecules. *J. Comput Chem.* **33**, 1637–1642 (2012).
9. Cioslowski, J. & Strasburger, K. Symmetry equiincidence of natural orbitals. *J. Phys. Chem. Lett.* **14**, 9296–9303 (2023).
10. Heit, Y. N., Gendron, F. & Autschbach, J. Calculation of dipole-forbidden 5f absorption spectra of uranium(V) hexa-halide complexes. *J. Phys. Chem. Lett.* **9**, 887–894 (2018).
11. Sinha, N. & Wenger, O. S. Photoactive metal-to-ligand charge transfer excited states in $3d^6$ complexes with $Cr^0$, $Mn^I$, $Fe^{II}$, and $Co^{III}$. *J. Am. Chem. Soc.* **145**, 4903–4920 (2023).
12. Kazmierczak, N. P., Mirzoyan, R. & Hadt, R. G. The Impact of ligand field symmetry on molecular qubit coherence. *J. Am. Chem. Soc.* **143**, 17305–17315 (2021).
13. Kasha, M., et al. 50 YEARS OF THE JABLONSKI DIAGRAM. *Acta Phys. Pol. A* **71**, 661–670 (1987).
14. Griffiths, J. et al. *Colour and Constitution of Organic Molecules*. (Academic Press, 1976).
15. Lyu, Z.-Y., Dong, H., Yang, X.-F., Sun, L.-D. & Yan, C.-H. Highly polarized upconversion emissions from lanthanide-doped LiYF4 crystals as spatial orientation indicators. *J. Phys. Chem. Lett.* **12**, 11288–11294 (2021).
16. Chen, X. et al. Achieving ultralong room-temperature phosphorescence in two-dimensional metal halide perovskites by alkyl chain engineering. *J. Phys. Chem. Lett.* **14**, 8638–8647 (2023).
17. Zhao, J. et al. Circularly polarized luminescence from achiral single crystals of hybrid manganese halides. *J. Am. Chem. Soc.* **141**, 15755–15760 (2019).
18. Rajca, A., Safronov, A., Rajca, S. & Wongsriratanakul, J. $D_2$-symmetric dimer of 1,1′-binaphthyl and its chiral π-conjugated carbodianion. *J. Am. Chem. Soc.* **122**, 3351–3357 (2000).
19. Ananias, D., Paz, F. A. A., Yufit, D. S., Carlos, L. D. & Rocha, J. Photoluminescent thermometer based on a phase-transition lanthanide silicate with unusual structural disorder. *J. Am. Chem. Soc.* **137**, 3051–3058 (2015).
20. Rudowicz, C. & Bramley, R. On standardization of the spin Hamiltonian and the ligand field Hamiltonian for orthorhombic symmetry. *J. Chem. Phys.* **83**, 5192–5197 (1985).
21. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
22. Cruz-Cabeza, A. J. & Bernstein, J. Conformational polymorphism. *Chem. Rev.* **114**, 2170–2191 (2014).
23. Muller, P. Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994). *Pure Appl. Chem.* **66**, 1077–1184 (1994).
24. Waterbeemd, H. V. D. et al. Glossary of terms used in computational drug design (IUPAC Recommendations 1997. *Pure Appl. Chem.* **69**, 1137–1152 (1997).
25. Hartshorn, R. M., Hellwich, K.-H., Yerin, A., Damhus, T. & Hutton, A. T. Brief guide to the nomenclature of inorganic chemistry. *Pure Appl. Chem.* **87**, 1039–1049 (2015).
26. Maheshwary, S., Patel, N., Sathyamurthy, N., Kulkarni, A. D. & Gadre, S. R. Structure and stability of water clusters $(H_2O)n$, n = 8–20: an ab initio investigation. *J. Phys. Chem. A* **105**, 10525–10537 (2001).
27. Jensen, C. V., Vogt, E., Poulsen, A. S. & Kjaergaard, H. G. Room temperature gas phase equilibrium constants of the methanol dimer, trimer, and tetramer. *J. Phys. Chem. A* **128**, 6382–6391 (2024).
28. Wernet, P. et al. The structure of the first coordination shell in liquid water. *Science* **304**, 995–999 (2004).
29. Eaves, J. D. et al. Hydrogen bonds in liquid water are broken only fleetingly. *Proc. Natl Acad. Sci. USA* **102**, 13019–13022 (2005).
30. Gruenloh, C. J. et al. Infrared spectrum of a molecular ice cube: The S-4 and D-2d water octamers in benzene-(water)(8). *Science* **276**, 1678–1681 (1997).
31. Smith, J. D. et al. Energetics of hydrogen bond network rearrangements in liquid water. *Science* **306**, 851–853 (2004).
32. Malenkov, G. Liquid water and ices: understanding the structure and physical properties. *J. Phys.-Condens. Mat.* https://doi.org/10.1088/0953-8984/21/28/283101 (2009).
33. Myneni, S. et al. Spectroscopic probing of local hydrogen-bonding structures in liquid water. *J. Phys.-Condens Mat.* **14**, L213–L219 (2002).
34. Rudberg, E., Rubensson, E. H., Sałek, P. & Kruchinina, A. Ergo: an open-source program for linear-scaling electronic structure calculations. *SoftwareX* **7**, 107–111 (2018).
35. Algara-Siller, G. et al. Square ice in graphene nanocapillaries. *Nature* **519**, 443 (2015).
36. Takaiwa, D., Hatano, I., Koga, K. & Tanaka, H. Phase diagram of water in carbon nanotubes. *Proc. Natl Acad. Sci. USA* **105**, 39–43 (2008).
37. Pauling, L. The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *J. Am. Chem. Soc.* **57**, 2680–2684 (1935).
38. Kuhs, W. F., Finney, J. L., Vettier, C. & Bliss, D. V. Structure and hydrogen ordering in ice-vi, ice-vii, and ice-viii by neutron powder diffraction. *J. Chem. Phys.* **81**, 3612–3623 (1984).
39. Mishima, O. & Stanley, H. E. The relationship between liquid, supercooled and glassy water. *Nature* **396**, 329–335 (1998).
40. Wang, C.-R., Sugai, T., Kai, T., Tomiyama, T. & Shinohara, H. Production and isolation of an ellipsoidal C80 fullerene. *Chem. Commun.*, **2000**, 557–558 (2000).
41. Bethune, D. S. et al. Vibrational raman and infrared-spectra of chromatographically separated C60 and C70 fullerene clusters. *Chem. Phys. Lett.* **179**, 181–186 (1991).
42. Iwasa, Y. et al. NEW phases of C-60 synthesized at high-pressure. *Science* **264**, 1570–1572 (1994).
43. Thess, A. et al. Crystalline ropes of metallic carbon nanotubes. *Science* **273**, 483–487 (1996).

44. Casanova, D. & Alemany, P. Quantifying the symmetry content of the electronic structure of molecules: molecular orbitals and the wave function. *Phys. Chem. Chem. Phys.* **12**, 15523–15529 (2010).
45. Boltalina, O. V., Borschevskii, A. Y., Sidorov, L. N., Street, J. M. & Taylor, R. Preparation of C60F36 and C70F36/38/40. *Chem. Commun.* https://doi.org/10.1039/cc9960000529 (1996).
46. Maki, G., et al. Ligand field theory of Ni(II) complexes. I. electronic energies and singlet ground-state conditions of Ni(II) complexes of different symmetries. *J. Chem. Phys.* **28**, 651–662 (1958).
47. Mouesca, J. M., Chen, J. L., Noodleman, L., Bashford, D. & Case, D. A. Density-functional poisson-boltzmann calculations of redox potentials for iron-sulfur clusters. *J. Am. Chem. Soc.* **116**, 11898–11914 (1994).
48. Murray-Rust, P., Burgi, H. B. & Dunitz, J. D. Distortions of MX4 molecules from Td symmetry. II. Analysis of PO4, SO4 and AlCl4 species. *Acta Crystallogr. Sect. B* **34**, 1793–1803 (1978).
49. Gewirth, A. A. & Solomon, E. I. Electronic-structure of plastocyanin - excited-state spectral features. *J. Am. Chem. Soc.* **110**, 3811–3819 (1988).
50. Waldmann, O. et al. Relation between electrostatic charge density and spin hamiltonian models of ligand field in lanthanide complexes. *Inorg. Chem.* https://doi.org/10.1021/acs.inorgchem.4c04392 (2025).
51. Rao, S. V., Maganas, D., Sivalingam, K., Atanasov, M. & Neese, F. Extended active space ab initio ligand field theory: applications to transition-metal ions. *Inorg. Chem.* **63**, 24672–24684 (2024).
52. Leyser da Costa Gouveia, T., Maganas, D. & Neese, F. Restricted open-shell hartree–fock method for a general configuration state function featuring arbitrarily complex spin-couplings. *J. Phys. Chem. A* **128**, 5041–5053 (2024).
53. Aakesson, R., Pettersson, L. G. M., Sandstroem, M. & Wahlgren, U. Ligand Field Effects in the Hydrated Divalent and Trivalent Metal Ions of the First and Second Transition Periods. *J. Am. Chem. Soc.* **116**, 8691–8704 (2002).
54. Tsukahara, N. et al. Adsorption-induced switching of magnetic anisotropy in a single iron(II) phthalocyanine molecule on an oxidized Cu(110) surface. *Phys. Rev. Lett.* **102**, 167203 (2009).
55. Janiak, C. et al. Engineering coordination polymers towards applications. *Dalton transactions*, **2003**, 2781–2804, (2003).
56. Bousseksou, A., Molnár, G., Salmon, L. & Nicolazzi, W. Molecular spin crossover phenomenon: recent achievements and prospects. *Chem. Soc. Rev.* **40**, 3313–3335 (2011).
57. Ogawa, T., Sinha, N., Pfund, B., Prescimone, A. & Wenger, O. S. Molecular design principles to elongate the metal-to-ligand charge transfer excited-state lifetimes of square-planar Nickel(II) complexes. *Journal of the American Chemical Society*, https://doi.org/10.1021/jacs.2c08838 (2022).
58. Oppermann, M., Zinna, F., Lacour, J. & Chergui, M. Chiral control of spin-crossover dynamics in Fe (II) complexes. *Nat. Chem.* **14**, 739 (2022).
59. Bramwell, S. T. & Gingras, M. J. Spin ice state in frustrated magnetic pyrochlore materials. *Science* **294**, 1495–1501 (2001).
60. Stöhr, J. Exploring the microscopic origin of magnetic anisotropies with X-ray magnetic circular dichroism (XMCD) spectroscopy. *J. Magn. Magn. Mater.* **200**, 470–497 (1999).
61. Groot, F. d. Multiplet effects in X-ray spectroscopy. *Coord. Chem. Rev.* **249**, 31–63 (2005).
62. Long, J., Guari, Y., Ferreira, R. A. S., Carlos, L. D. & Larionova, J. Recent advances in luminescent lanthanide based Single-Molecule Magnets. *Coord. Chem. Rev.* **363**, 57–70 (2018).
63. Bogani, L. & Wernsdorfer, W. Molecular spintronics using single-molecule magnets. *Nat. Mater.* **7**, 179–186 (2008).
64. Rinehart, J. D. & Long, J. R. Exploiting single-ion anisotropy in the design of f-element single-molecule magnets. *Chem. Sci.* **2**, 2078–2085 (2011).
65. Goodwin, C. A. P., Ortu, F., Reta, D., Chilton, N. F. & Mills, D. P. Molecular magnetic hysteresis at 60 kelvin in dysprosocenium. *Nature* **548**, 439–442 (2017).
66. Li, H. D., Wu, S. G. & Tong, M. L. Lanthanide-radical single-molecule magnets: current status and future challenges. *Chem. Commun.* **59**, 6159–6170 (2023).
67. Long, J. R., Chilton, N. F., Vincent, A. H. & Whyatt, Y. L. Strong axiality in a dysprosium(III) bis(borolide) complex leads to magnetic blocking at 65 K. *J. Am. Chem. Soc.* **145**, 1572–1579 (2023).
68. Meihaus, K. R. & Long, J. R. Magnetic blocking at 10 K and a dipolar-mediated avalanche in salts of the bis(eta8-cyclooctatetraenide) complex [Er(COT)2]. *J. Am. Chem. Soc.* **135**, 17952–17957 (2013).
69. Liu, J. L., Chen, Y. C. & Tong, M. L. Symmetry strategies for high performance lanthanide-based single-molecule magnets. *Chem. Soc. Rev.* **47**, 2431–2453 (2018).
70. Randall McClain, K. et al. High-temperature magnetic blocking and magneto-structural correlations in a series of dysprosium(III) metallocenium single-molecule magnets. *Chem. Sci.* **9**, 8492–8503 (2018).
71. Chilton, N. F., Goodwin, C. A., Mills, D. P. & Winpenny, R. E. The first near-linear bis(amide) f-block complex: a blueprint for a high temperature single molecule magnet. *Chem. Commun. (Camb.)* **51**, 101–103 (2015).
72. Liu, J.-L. et al. Switching the anisotropy barrier of a single-ion magnet by symmetry change from quasi-D5h to quasi-Oh. *Chem. Sci.* https://doi.org/10.1039/c3sc50843a (2013).
73. Aldamen, M. A., Clemente-Juan, J. M., Coronado, E., Marti-Gastaldo, C. & Gaita-Arino, A. Mononuclear lanthanide single-molecule magnets based on polyoxometalates. *J. Am. Chem. Soc.* **130**, 8874–8875 (2008).
74. Le Roy, J. J., Korobkov, I. & Murugesu, M. A sandwich complex with axial symmetry for harnessing the anisotropy in a prolate erbium(III) ion. *Chem. Commun. (Camb.* **50**, 1602–1604 (2014).
75. Ferguson, A. et al. A mixed-valence Co7 single-molecule magnet with C3 symmetry. *Chem. Commun. (Camb)*, **33**, 3473–3475 (2007).
76. Canaj, A. B. et al. Insight into D(6h) symmetry: targeting strong axiality in stable dysprosium(III) hexagonal bipyramidal single-ion magnets. *Angew. Chem. Int Ed. Engl.* **58**, 14146–14151 (2019).
77. Sun, W. B. et al. High symmetry or low symmetry, that is the question - high performance Dy(III) single-ion magnets by electrostatic potential design. *Chem. Sci.* **7**, 684–691 (2016).
78. Pedersen, K. S. et al. Design of single-molecule magnets: insufficiency of the anisotropy barrier as the sole criterion. *Inorg. Chem.* **54**, 7600–7606 (2015).
79. Kofod, N., Nawrocki, P., Platas-Iglesias, C. & Sørensen, T. J. Electronic structure of ytterbium(III) solvates—a combined spectroscopic and theoretical study. *Inorg. Chem.* **60**, 7453–7464 (2021).
80. Prokop, K. A. et al. Unusual emission from 2F5/2 two lowest excited levels of Yb3+-doped LuPO4 nano/micro-crystalline materials. *J. Lumin.* **275**, 120777 (2024).
81. Goodwin, C. A. P., Reta, D., Ortu, F., Chilton, N. F. & Mills, D. P. Synthesis and electronic structures of heavy lanthanide metallocenium cations. *J. Am. Chem. Soc.* **139**, 18714–18724 (2017).
82. Binnemans, K. Interpretation of europium(III) spectra. *Coord. Chem. Rev.* **295**, 1–45 (2015).
83. Duan, C.-K. & Tanner, P. A. What use are crystal field parameters? a chemist's viewpoint. *J. Phys. Chem. A* **114**, 6055–6062 (2010).
84. Nawrocki, P. R. & Sørensen, T. J. Optical spectroscopy as a tool for studying the solution chemistry of neodymium(iii). *Phys. Chem. Chem. Phys.* **25**, 19300–19336 (2023).
85. Nielsen, V. R. M., Grasser, M., Mortensen, S. S., Le Guennic, B. & Sørensen, T. J. Electronic structure of a neodymium(III)

tris(oxidiacetate) complex from luminescence data and Ab initio calculations. *Inorg. Chem.* **63**, 18596–18607 (2024).
86. Zabrodsky, H., Peleg, S. & Avnir, D. Continuous symmetry measures. 2. Symmetry groups and the tetrahedron. *J. Am. Chem. Soc.* **115**, 8278–8289 (1993).
87. Dryzun, C., Alemany, P., Casanova, D. & Avnir, D. A continuous symmetry analysis of chemical bonding. *Chemistry* **17**, 6129–6141 (2011).
88. Alemany, P., Casanova, D. & Dryzun, C. Electronic structure and symmetry in conjugated π-electron systems. *Chem. – A Eur. J.* **17**, 14896–14906 (2011).
89. Vonci, M. et al. Rationalization of anomalous pseudocontact shifts and their solvent dependence in a series of C3-symmetric lanthanide complexes. *J. Am. Chem. Soc.* **139**, 14166–14172 (2017).
90. Zabrodsky, H., Peleg, S. & Avnir, D. Continuous symmetry measures. *J. Am. Chem. Soc.* **114**, 7843–7851 (1992).
91. Tuvi-Arad, I., Shalit, Y. & Alon, G. CSM software: continuous symmetry and chirality measures for quantitative structural analysis. *J. Chem. Inf. Modeling* **64**, 5375–5380 (2024).
92. Nawrocki, P. R., Kofod, N., Juelsholt, M., Jensen, K. M. Ø & Sørensen, T. J. The effect of weighted averages when determining the speciation and structure–property relationships of europium(iii) dipicolinate complexes. *Phys. Chem. Chem. Phys.* **22**, 12794–12805 (2020).
93. Nielsen, L. G. & Sørensen, T. J. Including and declaring structural fluctuations in the study of lanthanide(iii) coordination chemistry in solution. *Inorg. Chem.* **59**, 94–105 (2019).
94. Pinsky, M. & Avnir, D. Continuous symmetry measures. 5. The classical polyhedra. *Inorg. Chem.* **37**, 5575–5582 (1998).
95. Pinsky, M., Dryzun, C., Casanova, D., Alemany, P. & Avnir, D. Analytical methods for calculating Continuous Symmetry Measures and the Chirality Measure. *J. Comput. Chem.* **29**, 2712–2721 (2008).
96. Waroquiers, D. et al. ChemEnv: a fast and robust coordination environment identification tool. *Acta Crystallogr. Sect. B* **76**, 683–695 (2020).
97. Casanova, D. et al. Minimal distortion pathways in polyhedral rearrangements. *J. Am. Chem. Soc.* **126**, 1755–1763 (2004).
98. Alvarez, S. et al. Shape maps and polyhedral interconversion paths in transition metal chemistry. *Coord. Chem. Rev.* **249**, 1693–1708 (2005).
99. Ruiz-Martínez, A. & Alvarez, S. Stereochemistry of compounds with coordination number ten. *Chem. – A Eur. J.* **15**, 7470–7480 (2009).
100. Ruiz-Martínez, A., Casanova, D. & Alvarez, S. Polyhedral structures with an odd number of vertices: nine-coordinate metal compounds. *Chem. – A Eur. J.* **14**, 1291–1303 (2008).
101. Zabrodsky, H., Peleg, S. & Avnir, D. Symmetry as a continuous feature. *IEEE Trans. pattern Anal. Mach. Intell.* **17**, 1154–1166 (1995).
102. Zabrodsky, H. & Avnir, D. Continuous symmetry measures. 4. Chirality. *J. Am. Chem. Soc.* **117**, 462–473 (1995).
103. Llunell, M., Casanova, D., Cirera, J., Alemany, P. & Alvarez, S. SHAPE. *version 2.1. Universitat de Barcelona, Barcelona, Spain 2103*, (2013).
104. Keinan, S. & Avnir, D. Continuous symmetry analysis of tetrahedral/planar distortions. Copper chlorides and other AB(4) species. *Inorg. Chem.* **40**, 318–323 (2001).
105. Pinsky, M., Lipkowitz, K. B. & Avnir, D. Continuous symmetry measures. VI. the relations between polyhedral point-group/subgroup symmetries. *J. Math. Chem.* **30**, 109–120 (2001).
106. Alvarez, S., Avnir, D., Llunell, M. & Pinsky, M. Continuous symmetry maps and shape classification. the case of six-coordinated metal compounds. *N. J. Chem.* **26**, 996–1009 (2002).
107. Alvarez, S., Alemany, P. & Avnir, D. Continuous chirality measures in transition metal chemistry. *Chem. Soc. Rev.* **34**, 313–326 (2005).
108. Dryzun, C. Continuous symmetry measures for complex symmetry group. *J. Comput. Chem.* **35**, 748–755 (2014).
109. Alemany, P., Casanova, D., Alvarez, S., Dryzun, C. & Avnir, D. Continuous Symmetry Measures: A New Tool in Quantum Chemistry. In Reviews in Computational Chemistry (eds Parrill, A. L. & Lipkowitz, K. B.). https://doi.org/10.1002/9781119356059.ch7 (2017).
110. Cirera, J., Ruiz, E. & Alvarez, S. Continuous shape measures as a stereochemical tool in organometallic chemistry. *Organometallics* **24**, 1556–1562 (2005).
111. Thomsen, M. S., Anker, A. S., Kacenauskaite, L. & Sørensen, T. J. We are never ever getting (back to) ideal symmetry: structure and luminescence in a ten-coordinated europium(III) sulfate crystal. *Dalton Trans.* https://doi.org/10.1039/D2DT01522F (2022).
112. Cirera, J., Ruiz, E. & Alvarez, S. Stereochemistry and spin state in four-coordinate transition metal compounds. *Inorg. Chem.* **47**, 2871–2889 (2008).
113. Cirera, J., Alemany, P. & Alvarez, S. Mapping the stereochemistry and symmetry of tetracoordinate transition-metal complexes. *Chemistry* **10**, 190–207 (2004).
114. Casanova, D., Alemany, P., Bofill, J. M. & Alvarez, S. Shape and symmetry of heptacoordinate transition-metal complexes: structural trends. *Chem. – A Eur. J.* **9**, 1281–1295 (2003).
115. Casanova, D., Llunell, M., Alemany, P. & Alvarez, S. The rich stereochemistry of eight-vertex polyhedra: a continuous shape measures study. *Chemistry* **11**, 1479–1494 (2005).
116. Ruiz-Martinez, A., Casanova, D. & Alvarez, S. Polyhedral structures with an odd number of vertices: nine-atom clusters and supramolecular architectures. *Dalton Trans*, **2008**, 2583–2591, (2008).
117. Link, L. & Niewa, R. Polynator: a tool to identify and quantitatively evaluate polyhedra and other shapes in crystal structures. *J. Appl. Crystallogr.* **56**, 1855–1864 (2023).
118. Cumby, J. & Attfield, J. P. Ellipsoidal analysis of coordination polyhedra. *Nat. Commun.* **8**, 14235 (2017).
119. Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A* **32**, 922–923 (1976).
120. Pinsky, M. et al. Symmetry operation measures. *J. Computational Chem.* **29**, 190–197 (2008).
121. Beruski, O. & Vidal, L. N. Algorithms for computer detection of symmetry elements in molecular systems. *J. Comput Chem.* **35**, 290–299 (2014).
122. Casanova, D., Alemany, P. & Alvarez, S. Symmetry measures of the electron density. *J. Comput Chem.* **31**, 2389–2404 (2010).
123. Krumsieck, J. & Bröring, M. PorphyStruct: A Digital Tool for the Quantitative Assignment of Non-Planar Distortion Modes in Four-Membered Porphyrinoids. *Chem. – A Eur. J.* **27**, 11580–11588 (2021).
124. Kingsbury, C. J. & Senge, M. O. Quantifying near-symmetric molecular distortion using symmetry-coordinate structural decomposition. *Chem. Sci.* **15**, 13638–13649 (2024).
125. Collins, A., Cooper, R. I. & Watkin, D. J. Structure matching: measures of similarity and pseudosymmetry. *J. Appl. Crystallogr.* **39**, 842–849 (2006).
126. Baggio, R. A simple graphical method to pinpoint local pseudosymmetries in Z′ > 1 cases. *Acta Crystallogr. Sect. C.* **75**, 837–850 (2019).
127. Campbell, B. J., Stokes, H. T., Averett, T. B., Machlus, S. & Yost, C. J. The ISOTILT software for discovering cooperative rigid-unit rotations in networks of interconnected rigid units. *J. Appl. Crystallogr.* **54**, 1847–1856 (2021).
128. Brock, C. P. Approximate symmetry in low-symmetry space groups: P2 and C2. *Cryst. Growth Des.* **24**, 6211–6217 (2024).

129. Costamagna, M., Foscato, M., Grellscheid, D. & Jensen, V. R. Hypershape recognition: a general framework for moment-based molecular similarity. *J. Chem. Inf. Modeling* **65**, 5960–5972 (2025).
130. McFarlan, R. L. The Structure of Ice III. *J. Chem. Phys.* **4**, 253–259 (1936).
131. Goto, A., Hondoh, T. & Mae, S. The electron density distribution in ice Ih determined by single-crystal x-ray diffractometry. *J. Chem. Phys.* **93**, 1412–1417 (1990).
132. Xiao, S. et al. Molecular wires from contorted aromatic compounds. *Angew. Chem. Int. Ed.* **44**, 7390–7394 (2005).
133. Fedyanin, I. V. Control of supramolecular chirality in co-crystals of achiral molecules via stacking interactions and hydrogen bonding. *CrystEngComm* **24**, 2591–2601 (2022).
134. Juvanteny, A. et al. Unraveling the mechanism of hydrogen atom transfer by a nickel-hypochlorite species and the influence of electronic effects. *Inorg. Chem.* **63**, 14325–14334 (2024).
135. Smith, N. E., Bernskoetter, W. H., Hazari, N. & Mercado, B. Q. Synthesis and catalytic activity of PNP-supported iron complexes with ancillary isonitrile ligands. *Organometallics* **36**, 3995–4004 (2017).
136. Han, J., Nishihara, S., Inoue, K. & Kurmoo, M. On the nature of the structural and magnetic phase transitions in the layered perovskite-like (CH3NH3)2[FeIICl4]. *Inorg. Chem.* **53**, 2068–2075 (2014).
137. Chábera, P. et al. A low-spin Fe(iii) complex with 100-ps ligand-to-metal charge transfer photoluminescence. *Nature* **543**, 695–699 (2017).
138. Nielsen, V. R. M., Grasser, M., Le Guennic, B. & Sørensen, T. J. Neodymium(III) aqua ion as a model system in Ab initio crystal field analysis beyond point charges and crystal field theory. *Inorg. Chem.* **64**, 3463–3475 (2025).
139. Miao, L.-P. et al. Direct observation of geometric and sliding ferroelectricity in an amphidynamic crystal. *Nat. Mater.* **21**, 1158–1164 (2022).
140. Yabuuchi, Y. et al. Geometric tuning of coordinatively unsaturated copper(i) sites in metal–organic frameworks for ambient-temperature hydrogen storage. *J. Am. Chem. Soc.* **146**, 22759–22776 (2024).
141. Miyake, S. & Ueda, R. On Phase Transformation of BaTiO3. *J. Phys. Soc. Jpn.* **2**, 93–97 (1947).
142. Kwei, G. H., Lawson, A. C., Billinge, S. J. L. & Cheong, S. W. Structures of the ferroelectric phases of barium titanate. *J. Phys. Chem.* **97**, 2368–2377 (1993).
143. Storm Thomsen, M., Madsen, A. O. & Sørensen, T. J. Crystal structure and optical properties of a two-sited $Eu^{III}$ compound: an $Eu^{III}$ ion coordinated by two [EuIII(DOTA)]- complexes (DOTA is 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetate). *Acta Crystallographica Section C* https://doi.org/10.1107/S2053229621005647 (2021).
144. Storm Thomsen, M. & Sørensen, T. J. Delicate, a study of the structural changes in ten-coordinated La(III), Ce(III), Pr(III), Nd(III), Sm(III) and Eu(III) sulfates. *Dalton Trans.* **51**, 8964–8974 (2022).
145. Chen, Y.-C. et al. Symmetry-supported magnetic blocking at 20 K in pentagonal bipyramidal Dy(III) single-ion magnets. *J. Am. Chem. Soc.* **138**, 2829–2837 (2016).
146. Ishikawa, N., Sugita, M. & Wernsdorfer, W. Nuclear spin driven quantum tunneling of magnetization in a new lanthanide single-molecule magnet: bis(phthalocyaninato)holmium anion. *J. Am. Chem. Soc.* **127**, 3650–3651 (2005).
147. Chatterjee, A., Maslen, E. N. & Watson, K. J. The effect of the lanthanoid contraction on the nonaaqualanthanoid(III) tris(trifluoromethanesulfonates). *Acta Crystallogr. Sect. B* **44**, 381–386 (1988).
148. Thomas, J. R., Giansiracusa, M. J., Mole, R. A. & Sulway, S. A. Increasing the Symmetry around Lanthanide Ions: The Effect on Single-Ion Magnet Behavior and Electronic Structure. *Cryst. Growth Des.* **24**, 573–583 (2024).
149. Seitz, M., Oliver, A. G. & Raymond, K. N. The lanthanide contraction revisited. *J. Am. Chem. Soc.* **129**, 11153–11160 (2007).
150. Raymond, K. N., Wellman, D. L., Sgarlata, C. & Hill, A. P. Curvature of the lanthanide contraction: an explanation. *Comptes Rendus Chim.* **13**, 849–852 (2010).
151. Ellis, R. J. et al. Straining" to separate the rare earths: how the lanthanide contraction impacts chelation by diglycolamide ligands. *Inorg. Chem.* **56**, 1152–1160 (2016).
152. Boyle, T. J. et al. Synthesis and characterization of solvated lanthanide tris(trimethylsilyl)siloxides. *Inorg. Chem.* **61**, 5048–5059 (2022).
153. Colliard, I. & Deblonde, G. J. P. Similar but different: structural and spectroscopic characterization of series of neodymium, europium, americium, and curium coordination complexes. *J. Am. Chem. Soc.* **147**, 14455–14467 (2025).
154. Bagio, S. et al. Redox, spectroscopic and magnetic properties of C3-symmetric rare earth complexes featuring atypical ortho-dioxolene binding. *Dalton Trans.* **54**, 5061–5074 (2025).
155. Desreux, J. F. Nuclear magnetic-resonance spectroscopy of lanthanide complexes with a tetraacetic tetraaza macrocycle - unusual conformation properties. *Inorg. Chem.* **19**, 1319–1324 (1980).
156. Loncin, M. F., Desreux, J. F. & Merciny, E. Coordination of lanthanides by two polyamino polycarboxylic macrocycles: formation of highly stable lanthanide complexes. *Inorg. Chem.* **25**, 2646–2648 (1986).
157. Aime, S., Botta, M. & Ermondi, G. NMR study of solution structures and dynamics of lantbanide(iii) complexes of DOTA. *Inorg. Chem.* **31**, 4291–4299 (1992).
158. Aime, S. et al. Conformational and coordination equilibria on DOTA complexes of lanthanide metal ions in aqueous solution studied by 1H-NMR spectroscopy. *Inorg. Chem.* **36**, 2059–2068 (1997).
159. Parker, D., Dickins, R. S., Puschmann, H., Crossland, C. & Howard, J. A. K. Being excited by lanthanide coordination complexes: aqua species, chirality, excited-state chemistry, and exchange dynamics. *Chem. Rev.* **102**, 1977–2010 (2002).
160. Caravan, P., Ellison, J. J., McMurry, T. J. & Lauffer, R. B. Gadolinium(III) chelates as MRI contrast agents: Structure, dynamics, and applications. *Chem. Rev.* **99**, 2293–2352 (1999).
161. Sørensen, T. J. & Faulkner, S. Multimetallic lanthanide complexes: using kinetic control to define complex multimetallic arrays. *Acc. Chem. Res* **51**, 2493–2501 (2018).
162. Benetollo, F., Bombieri, G., Calabi, L., Aime, S. & Botta, M. Structural variations across the lanthanide series of macrocyclic DOTA complexes: insights into the design of contrast agents for magnetic resonance imaging. *Inorg. Chem.* **42**, 148–157 (2003).
163. Briganti, M. et al. Covalency and magnetic anisotropy in lanthanide single molecule magnets: the DyDOTA archetype. *Chem. Sci.* **10**, 7233–7245 (2019).
164. Cucinotta, G. et al. Magnetic anisotropy in a dysprosium/DOTA single-molecule magnet: beyond simple magneto-structural correlations. *Angew. Chem. Int. Ed.* **51**, 1606–1610 (2012).
165. Görller-Walrand, C. & Binnemans, K. Chapter 155 rationalization of crystal-field parametrization. *Handb. Phys. Chem. Rare Earths* **23**, 121–283 (1996).
166. Le Guennic, B. et al. in *Handbook on the Physics and Chemistry of Rare Earths* (Elsevier, 2025).
167. Nielsen, V. R. M., Nawrocki, P. R. & Sørensen, T. J. Electronic structure of neodymium(III) and europium(III) resolved in solution using high-resolution optical spectroscopy and population analysis. *J. Phys. Chem. A*, https://doi.org/10.1021/acs.jpca.3c00233 (2023).

168. Huang, P. et al. Unraveling the electronic structures of neodymium in LiLuF4 nanocrystals for ratiometric temperature sensing. *Adv. Sci.* **6**, 1802282 (2019).
169. Koningstein, J. A. & Geusic, J. E. Energy levels and crystal-field calculations of neodymium in yttrium aluminum garnet. *Phys. Rev.* **136**, A711–A716 (1964).
170. Hanwell, M. D. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* **4**, 17 (2012).
171. Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, I. I. I. W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024–10035 (1992).
172. Dahlmann, M., Erker, G., Fröhlich, R. & Meyer, O. Structural dichotomy in single-component ziegler catalyst systems: characterization of Zr···F and Zr···C-bonded structural types of group 4 metallocene [C4H6–B(C6F5)3] betaines. *Organometallics* **19**, 2956–2967 (2000).
173. Catalano, V. J., Malwitz, M. A. & Etogo, A. O. Pyridine substituted N-heterocyclic carbene ligands as supports for Au(I)–Ag(I) interactions: formation of a chiral coordination polymer. *Inorg. Chem.* **43**, 5714–5724 (2004).
174. Yi, C. S. et al. Acid-promoted homogeneous hydrogenation of alkenes catalyzed by the ruthenium–hydride complex (PCy3) 2(CO)(Cl)RuH: evidence for the formation of 14-electron species from the selective entrapment of the phosphine ligand. *Organometallics* **19**, 2909–2915 (2000).
175. He, X., Noll, B. C., Beatty, A., Mulvey, R. E. & Henderson, K. W. Ketone deprotonation mediated by mono- and heterobimetallic alkali and alkaline earth metal amide bases: Structural Characterization of Potassium, Calcium, and Mixed Potassium–Calcium Enolates. *J. Am. Chem. Soc.* **126**, 7444–7445 (2004).
176. Marc et al. Solid state organisation of C60 by inclusion crystallisation with triptycenes†. *Chem. Commun.* **1999**, 1709–1710, (1999).
177. Nakano, K. et al. Mechanism of Selective and Unselective Enclathration by a Host Compound Possessing Open, Flexible Host Frameworks. *Eur. J. Org. Chem.* **2003**, 2428–2436 (2003).
178. Zhang, J. et al. Solvent accommodation: functionalities can be tailored through co-crystallization based on 1:1 coronene-F4TCNQ charge-transfer complex. *ACS Appl. Mater. Interfaces* **9**, 1183–1188 (2017).
179. Ikumura, Y., Habuka, Y. & Hori, A. Co-crystal structures and Hirshfeld surface analysis of mesitylene and/or p-xylene solvated pseudopolymorphs of fully fluorinated Pd(II) complex. *Polyhedron* **197**, 115035 (2021).
180. Nikolayenko, V. I., Heyns, A. & Barbour, L. J. Threading the needle: guest transport in a versatile 0D porous molecular crystal. *Chem. Commun.* **53**, 11306–11309 (2017).
181. Singh, M. & Chopra, D. Diversity in mechanical response in donor–acceptor coupled cocrystal stoichiomorphs based on pyrene and 1,8-dinitroanthraquinone systems. *Cryst. Growth Des.* **18**, 6670–6680 (2018).
182. Hu, P., Du, K., Wei, F., Jiang, H. & Kloc, C. Crystal growth, HOMO–LUMO engineering, and charge transfer degree in perylene-FxTCNQ (x = 1, 2, 4) organic charge transfer binary compounds. *Cryst. Growth Des.* **16**, 3019–3027 (2016).
183. Graf, K. I., Huang, A. J., Meihaus, K. R. & Long, J. R. Methylamine separations enabled by cooperative ligand insertion in copper–carboxylate metal–organic frameworks. *J. Am. Chem. Soc.* **146**, 23943–23954 (2024).
184. Liu, Y. et al. A heteroleptic ferrous complex with mesoionic bis(1,2,3-triazol-5-ylidene) ligands: taming the MLCT excited state of iron(II). *Chem. – A Eur. J.* **21**, 3628–3639 (2015).
185. Prakash, O. et al. A stable homoleptic organometallic iron(IV) complex. *Chem. – A Eur. J.* **26**, 12728–12732 (2020).
186. Luz, C. et al. Construction of a boron chain on a single metal by dehydrocoupling of borane ligands. *J. Am. Chem. Soc.* **146**, 23741–23751 (2024).
187. Liu, J. et al. A stable pentagonal bipyramidal Dy(III) single-ion magnet with a record magnetization reversal barrier over 1000 K. *J. Am. Chem. Soc.* **138**, 5441–5450 (2016).
188. Blois, L. et al. Unusually large ligand field splitting in anionic europium(iii) complexes induced by a small imidazolic counterion. *Inorg. Chem.* **63**, 16861–16871 (2024).
189. Bünzli, J.-C. & Eliseeva, S. in *Lanthanide Luminescence* Vol. 7 *Springer Series on Fluorescence* (eds Pekka Hänninen & Harri Härmä) **3**, 1–45 (Springer Berlin Heidelberg, 2011).
190. Colliard, I. & Deblonde, G. J. P. Polyoxometalate ligands reveal different coordination chemistries among lanthanides and heavy actinides. *JACS Au* **4**, 2503–2513 (2024).
191. Beck, N. B. et al. Two neptunium(III) mellitate coordination polymers: completing the series Np–Cf of trans-uranic An(III) mellitates. *Inorg. Chem.* **61**, 17730–17737 (2022).
192. Tsunoda, M. & Gabbaï, F. P. μ6-η2:η2:η2:η2:η2:η2 as a new bonding mode for benzene. *J. Am. Chem. Soc.* **122**, 8335–8336 (2000).
193. Yamaura, H., Yamamoto, K. & Murahashi, T. Selective dimerization of a trinuclear mixed-metal sandwich complex: construction of an axially chiral metal skeleton. *Chem. Commun.* **57**, 9120–9123 (2021).
194. Mokuolu, Q. F. et al. Early–late, mixed-metal compounds supported by amidophosphine ligands. *Dalton Transactions*, **13**, 1960–1970, (2004).
195. Hong, S. H., Day, M. W. & Grubbs, R. H. Decomposition of a Key Intermediate in Ruthenium-Catalyzed Olefin Metathesis Reactions. *J. Am. Chem. Soc.* **126**, 7414–7415 (2004).
196. Miniajluk, N., Bondzior, B., Stefańska, D. & Dereń, P. J. Eu3+ ions in the highly symmetrical octahedral site in Ba2MgWO6 double perovskite. *J. Alloy. Compd.* **802**, 190–195 (2019).
197. Hölsä, J. & Porcher, P. Free ion and crystal field parameters for REOCl:Eu3+. *J. Chem. Phys.* **75**, 2108–2117 (1981).
198. Pinto Marinho, E. et al. Neodymium(III) and europium(III) picrate complexes with 2-azacyclononanone: spectroscopy and structure. *Polyhedron* **16**, 3519–3523 (1997).
199. Storm Thomsen, M., Andersen, H. O. B. & Sørensen, T. J. Long story short: donor set symmetry in [Eu(DOTA)(H2O)]− crystals determines the electronic structure. *Dalton Trans.* **51**, 14118–14124 (2022).
200. Kang, J.-G. & 김택진. Luminescence and Crystal-Field Analysis of Europium and Terbium Complexes with Oxydiacetate and 1,10-Phenanthroline. *Bulletin of the Korean Chemical Society* **26**, 1057–1064, (2005).
201. Mortensen, S. S., Nielsen, V. R. M. & Sørensen, T. J. Contrasting impact of coordination polyhedra and site symmetry on the electronic energy levels in nine-coordinated Eu(III) and Sm(III) crystals structures determined from single crystal luminescence spectra. *Dalton Transactions*, https://doi.org/10.1039/D4DT00157E (2024).
202. Apostolidis, C. et al. Zur Elektronenstruktur hochsymmetrischer Verbindungen der f-Elemente. 35 [1] Kristall- und Molekülstrukturen von Tris(hydrotris(1-pyrazolyl)borato)-lanthanid(III) (LnTp3; Ln = La, Eu) sowie Elektronenstruktur von EuTp3. *Z. für anorganische und Allg. Chem.* **628**, 2013–2025 (2002).
203. Carlson, E. H. & Dieke, G. H. The state of the Nd3+ ion as derived from the absorption and fluorescence spectra of NdCl3 and their zeeman effects. *J. Chem. Phys.* **34**, 1602–1609 (1961).
204. Gruber, J. B. & Satten, R. A. Analysis of the electronic spectra of neodymium ethylsulfate. *J. Chem. Phys.* **39**, 1455–1463 (1963).
205. Satten, R. A. Analysis of the spectrum of the Nd+++ Ion in the bromate crystal. *J. Chem. Phys.* **21**, 637–648 (1953).

206. Starynowicz, P., Bukietyn´ska, K., Ryba-Romanowski, W. & Dominiak-Dzik, G. & Goł¸b, S. Spectroscopy and structure of complexes with LNN bonds (III). Structure and spectroscopy of ethylenediammonium octaazi dodineodymate. *Polyhedron* **13**, 1069–1075 (1994).
207. Henderson, J. R., Muramoto, M. & Gruber, J. B. Spectrum of Nd3+ in lanthanide oxide crystals. *J. Chem. Phys.* **46**, 2515–2520 (1967).
208. Guttenberger, C., Unrecht, B., Reddmann, H. & Amberger, H. D. Electronic structures of organometallic complexes of f elements.: LV. Parametric analysis of the crystal field splitting pattern of tris(bis(trimethylsilyl)methyl)neodymium(III). *Inorg. Chim. Acta* **348**, 165–172 (2003).

## Acknowledgements
We thank David Avnir and Inbal Tuvi-Arad for valuable input and discussions during the revision of the manuscript. We thank the Novo Nordisk Foundation, Carlsbergfondet, Villum Fonden, and the University of Copenhagen for funding.

## Author contributions
T.J.S. secured funding and supervised the project. V.R.M.N. performed the experimental work, developed the methods and code, and did the data analysis. All other aspects of the research, including conceptualization and manuscript preparation, were contributed equally by both authors.

## Competing interests
The author declares no competing interests.

## Additional information
**Supplementary information** The online version contains supplementary material available at ( https://doi.org/10.1038/s41467-025-66054-7).

**Correspondence** and requests for materials should be addressed to Thomas Just Sørensen.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.