Probing the Multiple Structures of Vaterite through Combined Computational and Experimental Raman Spectroscopy.

M. De La Pierre, R. Demichelis, U. Wehrmeister, D. E. Jacob, P. Raiteri, J. D. Gale, R. Orlando

Probing the Multiple Structures of Vaterite through Combined Computational and Experimental Raman Spectroscopy.

JOURNAL OF PHYSICAL CHEMISTRY. C, NANOMATERIALS AND INTERFACES (2014) 118

DOI: 10.1021/jp509958d

Probing the multiple structures of vaterite through combined computational and experimental Raman spectroscopy

Marco De La Pierre,[1,2] Raffaella Demichelis,[1,] Ursula Wehrmeister,[3] Dorrit E. Jacob,[4] Paolo Raiteri[1], Julian D. Gale[1], Roberto Orlando[2]*

[1] Nanochemistry Research Institute, Department of Chemistry, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

[2] Dipartimento di Chimica, Università degli Studi di Torino and NIS - Nanostructured Interfaces and Surfaces - Centre of Excellence, Via Giuria 7, 10125 Torino, Italy

[3] Institute of Geosciences, Johannes Gutenberg – Universität Mainz, Germany

[4] Department of Earth and Planetary Sciences, Macquarie University, Sydney, NSW, Australia

KEYWORDS

vaterite, biomineralization, structure determination, Raman spectroscopy, *Herdamania momus*, chirality, layer stacking

ABSTRACT

First-principles Raman spectra have been computed for several new vaterite structural models that have been recently proposed, and compared with spectra recorded on a set of biogenic, geological and synthetic samples. This set includes new measurements collected on *Herdamania momus* spicules (Great Barrier Reef, Queensland, Australia), which are known to have purity and crystallinity that are higher than for other biogenic samples. Overall, due to the close structural connection between the various models, the computed Raman spectra are found to be broadly similar. However, the spectra obtained for the two most stable models (monoclinic *C2* and trigonal $P3_221$, corresponding to two different polytypes of vaterite) exhibit features that are in excellent agreement with the experimental spectra, whereas the other theoretical structures show minor peaks that are not observed experimentally. When comparing the spectra for the two lowest energy structural models (*C2* and $P3_221$), the differences are too small to discriminate between these candidates. The Raman spectrum of *Herdamania momus* is of higher quality with respect to spectra obtained in previous studies on other biogenic samples. However, there is no significant and systematic difference with respect to samples of geological and synthetic origin.

INTRODUCTION

Vaterite is a metastable anhydrous calcium carbonate polymorph often found in nature as a result of biomineralization.[1-3] In particular, this mineral can appear either as a final polycrystalline product in the hard tissues of living organisms, or most frequently as an intermediate structure during calcium carbonate crystallization. [4]

Before the function of vaterite in the various biomineralization scenarios can be explained, ideally its structural features need to be fully understood. Of all of the numerous models

proposed during the past 60 years, Meyer's hypotheses (orthorhombic or hexagonal)[5,6] and Kamhi's hypothesis (hexagonal with partial occupancy of carbonate sites)[7] have been the most widely accepted and debated for decades, despite the fact that none of their structures can separately explain all spectroscopic and microscopic features that vaterite samples show.[8,9]

In the last five years major advances have been made in understanding the crystallographic structure of vaterite.[8,10-14] While providing a better, though not definitive, description of vaterite structural features, these studies have also raised a few issues that are still unresolved and may be crucial in understanding the mechanism of some of the various biomineralization pathways. For example, some remaining key questions include the following: How many polytypes of vaterite are there and why? What is their occurrence and why are their crystallites so small? Is vaterite actually chiral? Is it possible to obtain a single crystal, or do all samples contain multiple structures?

In 2009, Wang and Becker[10] first analyzed and compared the various crystallographic structures that had been proposed for vaterite up to that point, ruling out most of them and proposing a new hexagonal model, with $P6_522$ symmetry and an expanded unit cell containing 18 $CaCO_3$ units. This model was found to be more consistent with spectroscopic data relative to the other prior models, as it allows for independent CO_3^{2-} units within the unit cell. However, Raman spectra recorded on several synthetic, geological and biological samples presented evidence for at least 3 independent carbonate anions to be present in the unit cell, whereas the $P6_522$ model has only 2.[9] The earlier orthorhombic and $P6_3/mmc$ hexagonal models have also recently been ruled out by Balan *et al*.[15] in favor of a new hexagonal model, through simulation of the mixing energy resulting from sulphate substitutions into vaterite and re-analysis of the anomalous data presented by Fernandez-Diaz *et al*. in 2010.[16]

In 2012, two papers presenting seemingly different conclusions were published. In the first of these papers,[11] by some of the present authors, all of the main ordered structures available in the literature were re-examined by applying electronic structure calculations based on Density Functional Theory (DFT). This work showed that all of them are thermodynamically unstable, representing either a metastable phase (the orthorhombic structures with symmetry $P2_12_12_1$, resulting from relaxing Meyer's Pbnm model, and Ama2 are higher in energy than the most stable hexagonal* form $P3_221$ by +2.3 and +15.1 kJ/mol per formula unit, respectively), or a transition state ($P6_522$) between more stable hexagonal or pseudo-hexagonal arrangements ($P3_221$, $P6_5$, $P112_1$). It was also shown that inter-conversion between all of these crystalline phases should be possible at room temperature via rotation of the carbonate anions. In the second publication from 2012, Mugnaioli *et al*.[12] showed through automated diffraction tomography that their synthetic samples of vaterite have a monoclinic or triclinic structure belonging to a space group of either C2/c or C-1. In the same year, Wang and Becker[13] used molecular dynamics simulation to predict that vaterite could assume either a monoclinic C2/c structure or a hexagonal $P6_522$ structure.

A radical change in thinking with regard to the structure of this biomineral has occurred in 2013, as a result of both High-Resolution Transmission Electron Microscopy (HRTEM) on *Herdamania momus* spicules[8] and electronic structure calculations.[14] Kabalah-Amitai *et al*.[8] have shown for the first time that two distinct structures are present within the same sample of vaterite, one of which is hexagonal and the other remains of indeterminate symmetry. At the same time, Demichelis *et al*.[14] predicted from *ab initio* calculations that vaterite could exhibit

* We classify the structures derived from the $P6_522$ arrangement as "hexagonal" (though one is trigonal and one is monoclinic) based on the shape of their unit cell rather than on the crystal system.

multiple structures. Here, three almost isoenergetic families of structural configurations were found, with 6-layer hexagonal, 6-layer triclinic† and 2-layer monoclinic structures. The absolute minimum energy structures for each basin of configurations have $P3_221$, C1 (or P1) and C2 symmetry, respectively. Within these basins, which differ in terms of the stacking sequence of the carbonate layers only, carbonate rotations are also energetically possible at room temperature, which leads to an ensemble of structures that may interconvert. This is in agreement with the findings of Kabalah-Amitai *et al.*,[8] who hypothesized that the arrangement of calcium ions is conserved between their two structures, that the structures are extremely similar, and that they only differ in the relative orientations of the carbonate anions. Some of the possible theoretically predicted structures in the hexagonal basin are intrinsically chiral, and this is also the case for the C1 and C2 structures, though this property is hard to demonstrate with currently available experimental techniques, mainly due to the small size of vaterite crystallites.

The only major difference between the findings of Kabalah-Amitai *et al.*[8] and Demichelis *et al.*[14] is in regard to the symmetry and unit cell size of vaterite. In particular, the former authors find that their X-ray diffraction (XRD) pattern best fits with Kamhi's original model,[7] with a small hexagonal unit cell ($1/9^{th}$ of the cell employed by Demichelis *et al.*[14]) where carbonate sites have partial occupancy and an overall symmetry of $P6_3/mmc$. On the contrary, recent studies based on ^{43}Ca NMR spectroscopy and X-ray diffraction show that the only models that can simultaneously explain all the observed features are C2 and $P3_221$.[17]

---

† Despite the conventional choice for this space group being P1, C1 is used because it is obtained from the C-1 arrangement originally proposed by Mugnaioli *et al.*[12] This structure is triclinic, however the conventional cell has a quasi-monoclinic shape, with $\alpha \simeq \gamma \simeq 90^{\circ}$. To avoid further confusion on the structure of vaterite we have decided to keep the notation for the space group unchanged with respect to previous publications. However, we will refer to the C1 model as triclinic, since the primitive cell that is used for direct comparison with the other models is clearly triclinic.

In this paper we apply Raman spectroscopy, through combining results from computer simulations and experimental measurements, to understand whether this technique is able to provide further insights as to the complex structure of vaterite. We will focus here on the models resulting from Demichelis *et al*.,[14] leaving the delicate task of re-examining the $P6_3/mmc$ model for future investigations. Experimental data presented in Wehrmeister *et al*.[9,18] were re-examined in the light of the structural information from Kabalah-Amitai *et al*.[7] and Demichelis *et al*.[14] and new spectra were recorded. New samples include *Herdmania momus* spicules from the Great Barrier Reef (Queensland, Australia), the same kind of samples where Kabalah-Amitai *et al*. [8] identified two distinct structures.

COMPUTATIONAL METHODS

All quantum mechanical calculations were performed at the DFT level as implemented in CRYSTAL14.[19] The PBEsol functional[20] was used, together with an all-electron Gaussian type basis set already adopted in a number of studies on calcium carbonates, including recent Raman and IR investigations of the polymorphs calcite and aragonite.[21,22] The accuracy of this approach in describing the structure and properties of vaterite has already been documented in previous works.[11,14] Geometry optimization and the calculation of vibrational frequencies were performed following the same schemes and adopting the same parameters and thresholds as described in earlier publications.[11,14]

Raman vibrational spectra were computed for all of the seven minimum energy structures listed in Demichelis *et al*.[14] Raman intensities were obtained through an analytical approach based on combining gradients of mono- and bi-electronic integrals with a Coupled Perturbed Hartree-Fock/Kohn-Sham (CPHF/KS) scheme for the response of the crystalline orbitals to a

static dielectric field. Further details on this method can be found in Maschio *et al*.[23,24] Isotropic intensities for a polycrystalline sample were derived at the experimental conditions (temperature, laser frequency) according to Prosandeev *et al*.[25] Finally, simulated spectra were generated with inclusion of Lorentzian broadening. Being unable to compute broadening parameters (i.e. Full Width at Half Maximum, FWHM) within the adopted model, a single value of 5 cm^{-1} was set for all peaks. Although this is not a realistic value for the whole spectral range, it provides a reasonable resolution for the visual inspection of features in the computed spectra.

EXPERIMENTAL METHODS

All Raman spectra were recorded at room temperature using a HORIBA Jobin Yvon LabRAM HR (high resolution) 800 spectrometer equipped with a Si-based charge-coupled device (CCD) detector (Peltier-cooled), an integrated Olympus BX41 optical microscope and an automatized *x* – *y* stage. A 50× long-distance objective (numerical aperture 0.55) and a slit width of 100 μm were chosen. The scattered light was dispersed by a grating with 1800 grooves/mm. *Herdmania momus* spicules were measured using the 632.816 nm line of a HeNe laser for excitation with a laser spot size of *ca* 2 × 2 μm. All spectra were recorded twice. All other spectra were measured using the 532.21 nm line of a frequency-doubled Nd:YAG laser and details for these measurements are given in Wehrmeister *et al*.[9,18].

The wavenumber accuracy for both setups is ±0.5 cm^{-1} with a measured spectral resolution of 0.6 cm^{-1} (FWHM of the Rayleigh line). Spectra were calibrated using the 520.5 cm^{-1} band of a silicon wafer. Data acquisition and spectra treatment were carried out with the commercially available program LabSpec v4.02 (HORIBA Jobin Yvon GmbH). For all spectra, background subtraction and peak analysis were performed with the software OriginLabPro 7.5 equipped with

an additional peak-fitting module.

STRUCTURE AND SYMMETRY PROPERTIES

The structures of the seven models that will be used in this study for vaterite have already been extensively described in two previous publications. We present here a summary mostly focused on their symmetry features.

Vaterite can assume multiple structures belonging to at least three almost isoenergetic basins in the configuration space with different symmetries, namely triclinic, monoclinic and hexagonal (see Table 1). The main difference between these groups of structures is the stacking sequence of the CO_3^{2-} layers. Here the three different orientations of the plane of the carbonate groups within a layer are labeled as A, B and C, with a further ' (prime) symbol being used to distinguish between the two possible directions of the C-O bonds within this plane. In the hexagonal arrangement, the layers are in an AC'BB'CA' sequence, the unit cell has a hexagonal shape and contains 90 atoms. Structures belonging to the monoclinic group have a smaller unit cell (30 atoms) that contains 2 layers of carbonate anions. However, a supercell with hexagonal shape can be built, containing 90 atoms and 6 layers of carbonate units, arranged in an AA'BB'CC' sequence. The structures in the triclinic basin have a quasi-monoclinic conventional cell, with $\alpha \simeq ß \simeq 90^\circ$. However, for sake of comparison, it is more convenient to refer to the primitive cell, which contains 90 atoms and the sequence of CO_3^{2-} layers is A'AB'CA'A. The imbalance between the number of each of the six plane types leads to the distortion of the primitive cell away from a hexagonal lattice. Both triclinic structures have C1 symmetry and they will be referred to in the following sections as C1 (1) and C1 (2). A schematic representation of these structures can be found in Demichelis *et al*.[14], Figures 1, 2 and Table of Contents.

Within each basin, interconversion between several structures is possible through rotations of carbonate anions. In Demichelis *et al.*[14], we noted that some of the hexagonal structures belong to chiral space groups. In principle, most of the structures in the three basins are chiral, though for low symmetry structures (C1 and C2) both mirror images adopt the same space group.

The main structural features are summarized in Table 1, where the relative stability with respect to the most stable vaterite structure ($P3_221$) is also reported, and the relative population of each configuration has been recalculated in order to account for all the possible chiral images. This assumes that a Boltzmann equilibrium distribution exists between configurations, which may not be the case due to kinetic factors. The number of independent CO_3^{2-} units in the unit cell is also reported and ranges from 3 to 18.

Significantly, previous observations based on Raman spectroscopy have indicated that vaterite has at least 3 independent carbonate units, whereas most of the structures proposed in the literature, apart from the present ones, have no more than 2.[9] However, the presence of three or more independent carbonate units could also be due to the presence of multiple structures within the same sample, as highlighted by Kabalah-Amitai *et al.* [8]

THE COMPUTED RAMAN SPECTRA

The computed Raman spectra of vaterite for all of the hypothetical space groups considered here are reported in Figure 1. Due to the complexity of the vaterite structure, all spectra are very rich, including a large number of peaks with medium to high intensity. Furthermore, because of the close relationship of all of the candidate structures, most of the spectra appear quite similar at first sight. Therefore, it is convenient to divide the various regions as per the work of Wehrmeister *et al.*[9] and to analyze each region separately (Table 2). Expanded illustrations of the

different spectral regions are reported in Figure 2 and Figure 3, and Figures S1 and S2 in the Supplementary Information.

Several theoretical studies have demonstrated that hybrid functionals are a better choice for predicting vibrational properties of minerals[26-28]. However, we have decided to perform this study using the PBEsol functional for sake of consistency with our previous calculations on vaterite, where PBEsol was preferred for providing more realistic thermodynamic predictions. Deviations between computed (PBEsol) and experimental spectra for vaterite have been compared with those obtained for aragonite, which has a well defined structure and for which accurate Raman experiments are available, with peaks assigned and symmetry classified. For all significant regions of the spectrum, we have been able to estimate the systematic shift (reported in Table 2), as well as the error (3 cm^{-1}) based on standard deviation. Further details are available as Supplementary Information). It turns out that, at the PBEsol level, different regions of the Raman spectrum are affected by different systematic deviations from experiment. The computed spectra and frequencies reported here for vaterite have been corrected according to these estimated systematic shifts, and can be considered accurate enough for the purposes of this paper.

As a cross check, Raman spectra have been computed for some of the most stable structures using the hybrid B3LYP[29] approach. As expected, the spectral regions obtained with this functional are not affected by any significant systematic shift[21,22,30] with respect to the experiment. Overall, the comparison of the B3LYP spectra between the different structures and with the experiments leads to the same qualitative conclusions as obtained with the PBEsol spectra, once the latter have been corrected by the systematic shifts.

The symmetry classification of the Raman active modes according to the various space groups is reported in Table 3. The region containing the lattice modes, L, shows between 51 and 159

Raman active modes, which appear in the experimental spectra as 8 broad features. The other regions contain a smaller number of Raman active modes: 6 to 18 in the ν_1 and ν_2 regions (with 3 and 1 to 3 peaks usually identified in the experiments, respectively) and 12 to 36 modes in the ν_3 and ν_4 regions (2 to 6 experimental peaks for both).

In order to make the computed results comparable with the experimental findings, given the high number of modes in relatively small ranges of wavenumbers, in our analysis peaks have been assumed to be distinct when the difference between their ν_{max} values (see definition below) is at least 5 cm^{-1}. This definition is obviously somewhat arbitrary. However, repeating the classification with different threshold values gives very similar results, and does not change the main spectral features here discussed. Data for frequencies and intensities are reported in Table 4, Table 5 and in Table S3 in the Supplementary Information. Here, each peak is the result of the overlap of N_m computed modes, and has a frequency ν_{max} corresponding to the frequency of the most intense contributing mode, and an integrated intensity II that is the sum of the N_m intensities.

Regions ν_1 and ν_4

The regions containing CO_3^{2-} symmetric stretching (ν_1) and in-plane bending (ν_4) modes show the most intense and significant peaks, and are reported in Figure 2 and Figure 3 and in Table 4 and Table 5. There are three groups of spectra showing major similarities, especially in the ν_1 region; these arise from the structures with the following symmetries a) C2 and $P3_221$, b) Cc and $P6_5$, c) C1 structures and $P112_1$. Looking at the energetics, we note that structures belonging to these groups of spectra share very similar stabilities.

In the region of ν_1 all peaks are found in the range 1095-1069 cm^{-1} and integrated intensities are of the order of hundreds to thousands au (arbitrary units). Table 4 shows that all spectra have

features around 1082 cm^{-1} and 1070 cm^{-1} (columns 3 and 5, respectively). The former shows significantly lower intensity for C2, $P3_221$ and C1 (1) with respect to the other structures, whereas the latter has similar high intensities in all structures.

The configurations C2 and $P3_221$ both lack a peak around 1094 cm^{-1} (column 1 in Table 4) and a peak around 1079 cm^{-1} (column 4), with the relative intensity of the other peaks being very similar between the two structures. The Cc and $P6_5$ structures both lack the features around 1087 cm^{-1} (column 2) and around 1079 cm^{-1} (column 4), with the remaining peaks having roughly the same intensities. C1 and $P112_1$ structures have very similar features, apart from C1 (2) lacking the peak around 1079 cm^{-1} (column 4).

The ν_4 region is reported in Table 5 and in Figure 3. Peaks are found in two distinct sub-ranges, namely 769-740 and 687-661 cm^{-1}, and intensities are on average much lower than those associated with the ν_1 region, being typically of the order of a few hundred au. In the region 687-661 cm^{-1} (columns 13-16 in Table 5) the seven structures show very similar features. The main differences are one missing peak in structures Cc and $P6_5$ (column 14) and one missing peak for structure $P3_221$ (column 15). In this region, all structures show their most intense peak around 686 cm^{-1} (column 13).

The range between 769 and 740 cm^{-1} (columns 8-12) is more interesting since different structures show different behaviors. The C2 and $P3_221$ arrangements are the only ones lacking the peak at about 769-765 cm^{-1} (column 8) and having both peaks number 11 and 12; in both cases peak number 9 has an intensity 2-3 times larger than others. The Cc and $P6_5$ arrangements exhibit very similar features as in the previous cases, with two missing peaks with respect to the other structures. The C1 and $P112_1$ structures are the only ones having both peaks number 8 and

9, with the latter being twice as intense as the former. Note also that $P3_221$ is the only structure lacking peak number 10 at around 750 cm^{-1}.

Regions ν_2 and ν_3

Because of the low intensity, we will comment only briefly on these two regions of the spectrum. The range 900-860 cm^{-1} (Figure 3 and Table 4) hosts ν_2 out-of-plane bending modes of the carbonate subunits. All structures apart from C2 show only one peak in this region, at about 877-880 cm^{-1}. Structure C2 has two peaks with similar intensity at 881 and 875 cm^{-1}. The range 1540-1360 cm^{-1} corresponds to ν_3 asymmetric stretching modes of the carbonate anions (Figure S2 in the Supplementary Information). This region is characterized by a large number of modes having similar low intensities (this spectral feature is noisy or absent in the experiments) and we will not enter into deeper detail in the present discussion.

Lattice modes

An accurate description of the spectral features of lattice modes would be very challenging for two reasons. Firstly, the spectrum is very complex (up to 160 modes). Secondly, this region is very sensitive to impurities and, at the same time, it falls in a frequency range that is very close to the limit of the edge filter in the experiments, so that comparison with experimental spectra would not be possible. Therefore we will briefly discuss the region of the lattice modes, with the aim of outlining its most relevant features.

Spectra are shown in Figure S1 and the main peaks are listed in Table S3, respectively, in the Supplementary Information. We focus our analysis on two specific sub-ranges where the most intense peaks are found, namely 320-250 and 130-80 cm^{-1}. To further simplify the analysis, of all the peaks found within these two spectral ranges, Table S3 reports only those with integrated intensity larger than 1000 au. The most significant feature of the former region is that all but one

structures show a single peak that is far more intense than all others between 311 and 303 cm^{-1}. C2 has two intense peaks with nearly the same intensity, at 305 and 296 cm^{-1}. In the latter region all structures but one show their most intense peak between 115 and 101 cm^{-1}, being instead located at 92 cm^{-1} in the case of $P112_1$.

THE EXPERIMENTAL RAMAN SPECTRUM OF HERDMANIA MOMUS

Herdmania Momus[31] is a solitary ascidian found as a sessile filter feeder at shallow depths in the temperate and tropical oceans.[32] It is one of the few known organisms that form vaterite as a normal constituent of their endoskeleton. Each animal contains a large number of vaterite spicules in its body and tunic ranging in size between 100 μm and 2.5 mm.[33] Most recently, these vaterite spicules were studied in detail by Kabalah-Amitai *et al*.[8] Our study obtained new high-resolution Raman spectra on spicules from a different individual from the Great Barrier Reef, Australia, allowing for a consistent comparison with previous measurements on other vaterite samples.[9,18]

The Raman spectra measured on spicules from this *Herdmania Momus* sample is shown at the bottom of Figure 1. Expanded spectra for specific ranges are presented at the bottom of Figure 2 and Figure 3 and Figures S2-S3 in the Supplementary Information. Band positions and FWHM values extracted from these measurements are reported in Table 6. In general, the sample does not suffer from the luminescence issues encountered in previous studies on biogenic vaterite.[9,18] As a result, the spectrum exhibits an improved quality and more bands can be identified. It is worth noticing that the smaller FWHM and the absence of peaks corresponding to calcite and aragonite indicate that the sample analyzed in this study could be characterized by higher crystallinity and purity with respect to biogenic samples considered in the past.

Three bands can be identified in the ν_1 region, at 1075, 1081 and 1091 cm^{-1}, in line with previous measurements on other biogenic, geologic and synthetic vaterite samples.[9,18] The ν_4 region shows five bands at 673, 686, 739, 744 and 751 cm^{-1}, providing a richer picture compared to biogenic samples analyzed in a previous study,[9] where only 2 to 4 bands could be found. A similar study on other biogenic samples showed the presence of 6 peaks in this region,[18] as in geologic and synthetic vaterite.[9,18] The missing band is at 666-668 cm^{-1} and in the present measurements it could not be distinguished from the one at 673 cm^{-1}. However, note that this region is affected by a large signal-to-noise ratio.

In the lattice mode region, five bands are identified at 176, 209, 268, 302 and 336 cm^{-1}, which lie at similar frequencies compared to previous studies;[9,18] notably, the corresponding FWHM values are far smaller . The spectral region below 160 cm^{-1} has not been investigated because it is very close to the edge filter of the instrument. Finally, no peaks can be identified in the ν_2 and ν_3 regions because of the high signal–to-noise ratio. The latter regions exhibited similar problems also in previous studies: one band in region ν_2 was found only in one sample,[9] and the band in region ν_3 was observed only in Reference [18] through FT-Raman.

Overall, these measurements on *Herdmania Momus* exhibit the typical Raman bands found for vaterite samples investigated in the past and coming from different origins (synthetic, geologic, biogenic). Kabalah-Amitai *et al*.[8] found that at least two distinct structures were present in *Herdmania momus* samples, and computer simulation (see Section above) shows that the various possible vaterite structures exhibit very similar Raman features. Therefore, similarities between different measurements possibly imply that also other samples analyzed in the past may contain two (or even more) distinct structures.

COMPARISON BETWEEN COMPUTED AND EXPERIMENTAL SPECTRA

Let us focus on the ν_1 and ν_4 spectral regions, which are the ones where relevant spectral features can be identified. As mentioned in the previous sections, ν_2 has very low intensity in both simulation and experiment; ν_3 has either low intensity (simulation) or noisy broad bands or even no signal at all (experiment); the lattice region is very complex, sensitive to impurities and defects, and close to the limit of the instrument.

In region ν_1 (Table 4 and Table 6, Figure 2), two computed structures, namely C2 and $P3_221$, are able to provide a very satisfactory match to the experimental spectrum. They both have 3 peaks, with frequencies in good agreement with the 3 measured bands. Moreover, the distance between extreme peaks (Δ), *i.e.* between highest and lowest frequency peaks in the region, is 18 and 16 cm^{-1}, respectively, compared to the experimental value of 16 cm^{-1}. The other five structures all show a peak at around 1095-1093 cm^{-1} with no corresponding features in the measurements, which results in Δ values between 22 and 24 cm^{-1}.

Similar observations hold for the 769-740 cm^{-1} sub-range of ν_4 (Table 5 and Table 6, Figure 3), where computed spectra for C2 and $P3_221$ show again great similarity with the experiment. Here the 4 and 3 peaks, respectively, have frequencies in good agreement with the 3 measured bands; Δ is equal to 16 and 15 cm^{-1}, respectively, to be compared with an experimental value of 12 cm^{-1}. Also in this case, the remaining computed structures show an extra feature at 769-765 cm^{-1}, which raises Δ to 22-27 cm^{-1}. The second sub-range of this region, 690-650 cm^{-1}, has low intensity in the experiment and nothing that can be used to discriminate between structures in the calculations, such that no further information relevant to this discussion can be extracted in this region.

When trying to discriminate between C2 and $P3_221$, only minor differences are found, which can be summarized by stating that the former shows a larger number of minor peaks than the latter. The most remarkable example is probably the case of the (low intensity) ν_2 region, where C2 is the only computed structure to show two peaks, separated by 6 cm^{-1}, rather than a single peak. Experimental synthetic samples usually show a larger number of peaks than biogenic ones. However, it is hard to draw conclusions from this observation, as it is also possibly due to factors others than vaterite polytypism, such as the degree of crystallinity and signal-to-noise ratio.

Having computed the free energy differences between the seven structures, we were able to compute the Boltzmann probabilities for each of them (Table 1) and build Boltzmann averaged simulated spectra for the three crystallographic "basins" proposed by Demichelis *et al.*[14] However, the three spectra (not reported here) turn out to be very similar, and in all of them the presence of the extra peaks in ν_1 and ν_4 discussed above results in an imperfect agreement with the experimental spectrum.

Overall, the comparison indicates that both the two lowest energy structures proposed in Demichelis *et al.* [14], namely C2 and $P3_221$, are compatible with Raman measurements on vaterite. This is in strong agreement with a recent study showing that only these two structures are compatible with the observed ^{43}Ca NMR spectra, and that both structures are compatible with the recorded X-ray diffraction pattern.[17]

Simulated Raman spectra do not make it possible to discriminate as to whether the first, the second or both structures are present in the experimental samples. Nonetheless, C2 and $P3_221$ have hexagonal and monoclinic unit cells, respectively, which is compatible with the recent finding by Kabalah-Amitai *et al.*[8] of two distinct structures during TEM investigation of *Herdmania Momus*, one with a hexagonal cell and one still unidentified. The remaining five

proposed computer structures seem to be ruled out according to this study. This runs contrary to the room temperature intra-basin structural inter-conversion suggested by Demichelis *et al*. [14] on the basis of a CO_3^{2-} rotation mechanism, which would result in different domains being present in the same crystal, and thus in peaks from all the structures to appear in the spectra. A possibility is that the system becomes kinetically trapped in a subset of the possible stable structures. To verify this hypothesis, further investigation is required in order to determine the activation barriers for the interconversion between all of the different energy minima.

ASSOCIATED CONTENT

Supporting Information is available free of charge via the Internet at http://pubs.acs.org."

AUTHOR INFORMATION

Corresponding Author

* raffaella.demichelis@curtin.edu.au

Author Contributions

M.D.L.P., R.D., P.R. and J.D.G. have performed and analyzed the simulations. U.W. and D.E.J. have performed all measures and analyzed the experimental spectra. R.O. has provided a development version of the CRYSTAL code able to perform the calculation of Raman intensities, which M.D.L.P. and R.D. have generalized to include the adopted and other DFT functionals. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank the Australian Research Council for Funding through Discovery grant DP0986999 and Future Fellowships FT120100462 and FT130100463, and Curtin University for funding through the Curtin Research Fellowship scheme. iVEC facilities and the Australian National Computing Infrastructure are also acknowledged for the provision of computer time. We are grateful to Carmel McDougall (University of Queensland) for providing us with a suitable sample of *Herdmania momus*, and to David Quigley (University of Warwick) and Pupa Gilbert (University of Wisconsin-Madison) for fruitful discussions.

REFERENCES

(1) Hasse, B.; Ehrenberg, H.; Marxen, J. C.; Becker, W.; Epple, M. Calcium Carbonate Modifications in the Mineralized Shell of the Freshwater Snail Biomphalaria Glabrata. *Chemistry - A European Journal* **2000**, *6*, 3679–3685.

(2) Freeman, C. L.; Harding, J. H.; Quigley, D.; Rodger, P. M. Structural Control of Crystal Nuclei by an Eggshell Protein. *Angew Chem Int Ed* **2010**, *49*, 5135–5137.

(3) W, G. R.; Sharma, S. K.; Volk, E. Micro-Raman Spectral Study of Vaterite and Aragonite Otoliths of the Coho Salmon,Oncorhynchus kisutchR. *Comp. Biochem. Physiol.* **1997**, *118A*, 753–757.

(4) Rodriguez-Blanco, J. D.; Shaw, S.; Benning, L. G. The Kinetics and Mechanisms of Amorphous Calcium Carbonate (ACC) Crystallization to Calcite, via Vaterite. *Nanoscale* **2011**, *3*, 265–271.

(5) Meyer, H. J. Struktur Und Fehlordnung Des Vaterits. *Z Kristallogr* **1969**, *128*, 183–212.

(6) Meyer, H. J. Uber Vaterit Und Seine Struktur. *Angew Chem* **1959**, *71*, 678–679.

(7) Kamhi, S. R. On the Structure of Vaterite CaCO3. *Acta Cryst* **1963**, *16*, 770–772.

(8) Kabalah-Amitai, L.; Mayzel, B.; Kauffmann, Y.; Fitch, A. N.; Bloch, L.; Gilbert, P. U. P. A.; Pokroy, B. Vaterite Crystals Contain Two Interspersed Crystal Structures. *Science* **2013**, *340*, 454–457.

(9) Wehrmeister, U.; Soldati, A. L.; Jacob, D. E.; Häger, T.; Hofmeister, W. Raman Spectroscopy of Synthetic, Geological and Biological Vaterite: a Raman Spectroscopic Study. *J. Raman Spectrosc.* **2010**, *41*, 193–201.

(10) Wang, J.; Becker, U. Structure and Carbonate Orientation of Vaterite (CaCO3). *Amer Mineral* **2009**, *94*, 380–386.

(11) Demichelis, R.; Raiteri, P.; Gale, J. D.; Dovesi, R. A New Structural Model for Disorder in Vaterite From First-Principles Calculations. *CrystEngComm* **2012**, *14*, 44–47.

(12) Mugnaioli, E.; Andrusenko, I.; Schüler, T.; Loges, N.; Dinnebier, R. E.; Panthöfer, M.; Tremel, W.; Kolb, U. Ab Initio Structure Determination of Vaterite by Automated Electron Diffraction. *Angew Chem Int Ed* **2012**, *51*, 7041–7045.
(13) Wang, J.; Becker, U. Energetics and Kinetics of Carbonate Orientational Ordering in Vaterite Calcium Carbonate. *Amer Mineral* **2012**, *97*, 1427–1436.
(14) Demichelis, R.; Raiteri, P.; Gale, J. D.; Dovesi, R. The Multiple Structures of Vaterite. *Crystal Growth & Design* **2013**, *13*, 2247–2251.
(15) Balan, E.; Blanchard, M.; Pinilla, C.; Lazzeri, M. First-Principles Modeling of Sulfate Incorporation and 34S/32S Isotopic Fractionation in Different Calcium Carbonates. *Chemical Geology* **2014**, *374-375*, 84–91.
(16) Fernández-Díaz, L.; Fernández-González, Á.; Prieto, M. The Role of Sulfate Groups in Controlling CaCO3 Polymorphism. *Geochimica et Cosmochimica Acta* **2010**, *74*, 6064–6076.
(17) Burgess, K. M. N.; Bryce, D. L. On the Crystal Structure of the Vaterite Polymorph of CaCO3. a Calcium-43 Solid-State NMR and Computational Assessment. *Solid State Nuclear Magnetic Resonance* **2014,** doi: 10.1016/j.ssnmr.2014.08.003.
(18) Wehrmeister, U.; Jacob, D. E.; Soldati, A. L.; Loges, N.; Häger, T.; Hofmeister, W. Amorphous, Nanocrystalline and Crystalline Calcium Carbonates in Biological Materials. *J. Raman Spectrosc.* **2011**, *42*, 926–935.
(19) Dovesi, R.; Orlando, R.; Erba, A.; Zicovich-Wilson, C. M.; Civalleri, B.; Casassa, S.; Maschio, L.; Ferrabone, M.; De La Pierre, M.; D'Arco, P.; et al. CRYSTAL14: a Program for the Ab Initio Investigation of Crystalline Solids. *Int. J. Quantum Chem.* **2014**, *114*, 1287–1317.
(20) Perdew, J.; Ruzsinszky, A.; Csonka, G.; Vydrov, O.; Scuseria, G.; Constantin, L.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Phys. Rev. Lett.* **2008**, *100*, 136406.
(21) De La Pierre, M.; Carteret, C.; Maschio, L.; André, E.; Orlando, R.; Dovesi, R. The Raman Spectrum of CaCO3 Polymorphs Calcite and Aragonite: a Combined Experimental and Computational Study. *J. Chem. Phys.* **2014**, *140*, 164509.
(22) Carteret, C.; De La Pierre, M.; Dossot, M.; Pascale, F.; Erba, A.; Dovesi, R. The Vibrational Spectrum of CaCO3 Aragonite: a Combined Experimental and Quantum-Mechanical Investigation. *J. Chem. Phys.* **2013**, *138*, 014201.
(23) Maschio, L.; Kirtman, B.; Rérat, M.; Orlando, R.; Dovesi, R. Ab Initio Analytical Raman Intensities for Periodic Systems Through a Coupled Perturbed Hartree-Fock/Kohn-Sham Method in an Atomic Orbital Basis. I. Theory. *J. Chem. Phys.* **2013**, *139*, 164101.
(24) Maschio, L.; Kirtman, B.; Rérat, M.; Orlando, R.; Dovesi, R. Ab Initio Analytical Raman Intensities for Periodic Systems Through a Coupled Perturbed Hartree-Fock/Kohn-Sham Method in an Atomic Orbital Basis. II. Validation and Comparison with Experiments. *J. Chem. Phys.* **2013**, *139*, 164102.
(25) Prosandeev, S.; Waghmare, U.; Levin, I.; Maslar, J. First-Order Raman Spectra of AB1⁄2′B1⁄2″O3 Double Perovskites. *Phys. Rev. B* **2005**, *71*, 214307.
(26) De La Pierre, M.; Orlando, R.; Maschio, L.; Doll, K.; Ugliengo, P.; Dovesi, R. Performance of Six Functionals (LDA, PBE, PBESOL, B3LYP, PBE0, and WC1LYP) in the Simulation of Vibrational and Dielectric Properties of Crystalline Compounds. the Case of Forsterite Mg2SiO4. *J. Comput. Chem.* **2011**, *32*, 1775–

1784.
(27) Demichelis, R.; Civalleri, B.; Ferrabone, M.; Dovesi, R. On the Performance of Eleven DFT Functionals in the Description of the Vibrational Properties of Aluminosilicates. *Int. J. Quantum Chem.* **2010**, *110*, 406–415.
(28) Valenzano, L.; Torres, F. J.; Doll, K.; Pascale, F.; Zicovich-Wilson, C. M.; Dovesi, R. Ab InitioStudy of the Vibrational Spectrum and Related Properties of Crystalline Compounds; the Case of CaCO 3Calcite. *Zeitschrift für Physikalische Chemie* **2006**, *220*, 893–912.
(29) Becke, A. D. Density-Functional Thermochemistry. III. the Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
(30) Valenzano, L.; Torres, F. J.; Doll, K.; Pascale, F.; Zicovich-Wilson, C. M.; Dovesi, R. Ab InitioStudy of the Vibrational Spectrum and Related Properties of Crystalline Compounds; the Case of CaCO 3Calcite. *Zeitschrift für Physikalische Chemie* **2006**, *220*, 893–912.
(31) Savigny, M. J. C. Mémoires Sur Les Animaux Sans Vertèbres; Harvard College Library, 1816.
(32) Lowenstam, H. A.; Abbott, D. P. Vaterite: a Mineralization Product of the Hard Tissues of a Marine Organism (Ascidiacea). *Science* **1975**, *188*, 363–365.
(33) Lambert, G.; Lambert, C. C. Spicule Formation in the Solitary Ascidian, Herdmania Momus. *Journal of Morphology* **1987**, *192*, 145–159.

Table 1 Classification and symmetry of the seven structures of vaterite considered in the present study.

	SGN, HM	Chiral image	Crystal system	N_{op}	N_{at}	N_{CO3}	ΔU_E	ΔG	*p*
Triclinic	1, C1 (1)	C1	Triclinic[a]	1	90	18	0.4	0.1	0.087[d]
	1, C1 (2)	C1	Triclinic[a]	1	90	18	0.5	0.2	0.083[d]
Monoclinic	5, C2	C2	Monoclinic	2	30	4	0.1	0.4	0.077[d]
	9, Cc	None	Monoclinic	2	30	3	0.9	1.1	0.058
Hexagonal	170, $P6_5$	$P6_1$	Hexagonal	6	90	3	0.7	1.0	0.060[d]
	4, $P112_1$	$P112_1$	Monoclinic[b]	2	90	9	0.5	0.5	0.074[d]
	153, $P3_221$	$P3_121$	Trigonal[c]	6	90	4	0.0	0.0	0.090[d]

SGN: Space Group Number; HM: Hermann-Mauguin symbol; N_{op}: number of symmetry operators; N_{at}: number of atoms in the primitive cell; N_{CO3}: number of symmetry-independent CO_3^{2-} units in the unit cell; ΔU_E, ΔG: electronic energy and free energy differences at 298 K with respect to $P3_221$ [kJ/mol per formula unit], as calculated in Reference[14]; *p:* Boltzmann relative populations at 298 K recalculated to include all mirror images. [a] $\alpha \simeq \beta \simeq 90°$ [b] $a \simeq b$ and $\gamma \simeq 120°$ [c] $a=b$ and $\gamma=120°$ [d] Population of both mirror images.

Table 2 Nomenclature for the five regions of the vaterite spectrum and shifts applied to computed frequencies in order to correct for systematic shifts associated with the choice of Kohn-Sham Hamiltonian (estimated based on aragonite data; details available as Supplementary Information).

Symbol	Modes	Experimental Range [cm^{-1}]	Shift [cm^{-1}]
L	Lattice modes	0-340	-3.1
ν_4	In-plane bending of CO_3^{2-}	666-685 738-753	+24.3
ν_2	Out-of-plane bending of CO_3^{2-}	873-881	+44.4
ν_1	Symmetric stretching of CO_3^{2-}	1074-1091	+11.4
ν_3	Asymmetric stretching of CO_3^{2-}	1421-1555	+7.8

Table 3 Symmetry classification of Raman active modes in the seven hypothetical space groups.

	L	ν_4	ν_2	ν_1	ν_3
C1[a]	159 (A)	36 (A)	18 (A)	18 (A)	36 (A)
C2	51 (25A'+26B'')	12 (6A'+6B'')	6 (3A'+3B'')	6 (3A'+3B'')	12 (6A'+6B'')
Cc	51 (24A+27B)	12 (6A+6B)	6 (2A+4B)	6 (4A+2B)	12 (6A+6B)
$P6_5$	79 ($26E_1+27E_2+26A$)	18 ($6E_1+6E_2+6A$)	9 ($3E_1+3E_2+3A$)	9 ($3E_1+3E_2+3A$)	18 ($6E_1+6E_2+6A$)
$P112_1$	159 (80A+79B)	36 (18A+18B)	18 (9A+9B)	18 (9A+9B)	36 (18A+18B)
$P3_221$	78 ($53E+25A_1$)	18 ($12E+6A_1$)	8 ($6E+2A_1$)	10 ($6E+4A_1$)	18 ($12E+6A_1$)

Symbols for space groups in Table 1, regions of the spectrum in Table 2. [a] Both C1 structures share the same symmetry classification.

Table 4 Calculated Raman frequencies [cm^{-1}] and intensities [au, arbitrary units] of the seven structures considered here for vaterite, in the 1160-1020 and 900-860 cm^{-1} range (ν_1 and ν_2 modes).

		ν_1					ν_2	
		1	2	3	4	5	6	7
C1 (1)	ν_{max}	1093	1086	1083 *	1078	1071	878	
	II (N_m)	1701 (2)	4651 (7)	741 (4)	438 (1)	2074 (4)	138 (18)	
C1 (2)	ν_{max}	1094	1087	1083		1070	878	
	II (N_m)	1614 (2)	3842 (6)	2038 (6)		2112 (4)	99 (18)	
C2	ν_{max}		1087	1083 *		1069	881	875
	II (N_m)		7601 (3)	345 (2)		1615 (1)	26 (3)	21 (3)
Cc	ν_{max}	1094		1081		1072	878	
	II (N_m)	3780 (2)		2984 (2)		2838 (2)	221 (6)	
$P6_5$	ν_{max}	1093		1082		1070	877	
	II (N_m)	3836 (3)		2686 (3)		3096 (3)	88 (9)	
$P112_1$	ν_{max}	1095	1087	1083	1079	1072	878	
	II (N_m)	1357 (2)	3747 (6)	1430 (4)	891 (2)	2167 (4)	142 (18)	
$P3_221$	ν_{max}		1086	1082 *		1070	880	
	II (N_m)		7108 (6)	799 (2)		1672 (2)	40 (8)	

For each peak, ν_{max} is the frequency of the most intense contributing mode, II is the integrated intensity, i.e. the sum of the intensities of all the contributing modes, N_m is the number of contributing modes. Vertical alignment of the peaks is done so to simplify comparison among structures. Peaks marked with an asterisk are not distinct in the spectra, but yet have been reported, as their identification is particularly meaningful for the discussion.

Table 5 Calculated Raman frequencies and intensities of the seven structures considered here for vaterite, in the 780-640 cm^{-1} range (ν_4 modes).

		ν_4								
		8	9	10	11	12	13	14	15	16
C1 (1)	ν_{max}	768	756	750	744		686	680	671	665
	II (N_m)	179 (2)	434 (6)	333 (4)	469 (6)		355 (4)	66 (3)	45 (3)	166 (8)
C1 (2)	ν_{max}	766	758	751	744		685	680*	674	665
	II (N_m)	189 (2)	410 (5)	186 (3)	627 (8)		357 (4)	35 (2)	54 (3)	178 (9)
C2	ν_{max}		756	751	744	740	686	676	668	661
	II (N_m)		578 (2)	214 (1)	245 (2)	367 (1)	400 (2)	70 (1)	123 (2)	51 (1)
Cc	ν_{max}	765		751		740	684		669	665
	II (N_m)	558 (2)		340 (2)		524 (2)	377 (2)		119 (2)	123 (2)
$P6_5$	ν_{max}	768		750		743	685		671	664
	II (N_m)	488 (3)		360 (3)		590 (3)	384 (3)		110 (3)	124 (3)
$P112_1$	ν_{max}	769	754	749		742	686	679	673	664
	II (N_m)	186 (2)	516 (8)	377 (4)		337 (4)	333 (4)	53 (2)	77 (4)	160 (8)
$P3_221$	ν_{max}		758		746 *	743	687	679		666
	II (N_m)		777 (4)		236 (3)	402 (1)	335 (2)	127 (3)		172 (5)

For more details see caption to Table 4.

Table 6 Experimental Raman band positions and FWHM [cm^{-1}] of *Herdmania Momus* sample.

L	ν_4	ν_2	ν_1	ν_3
n.a. below 160	673 (12.2)	n.a.	1075 (4.4)	n.a.
176 (2.5)	686 (1.5)		1081 (6.2)	
209 (15.0)	739 (3.7)		1091 (5.0)	
268 (20.5)	744 (4.6)			
302 (18.5)	751 (6.8)			
336 (11.1)				

n.a., not analyzed because of large signal-to-noise ratio.

Figure 1 Calculated and experimental Raman spectra of vaterite. Computed spectra are for each of the seven structures considered here. Experimental spectrum measured on *Herdamania momus* spicules from Great Barrier Reef (Australia).

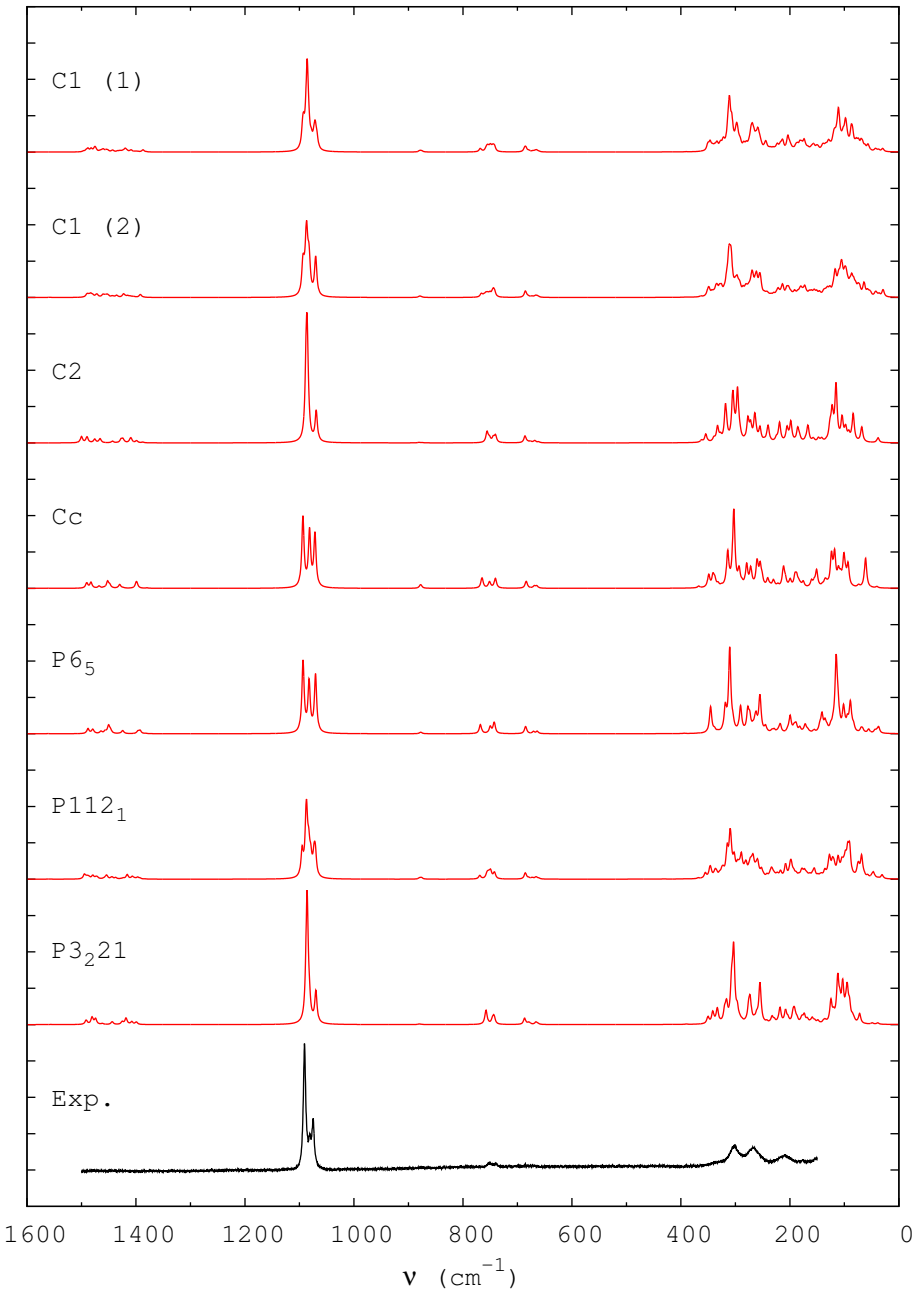

Figure 2 Calculated and experimental Raman spectra of vaterite: expansion of the 1160-1020 cm^{-1} range (ν_1 modes).

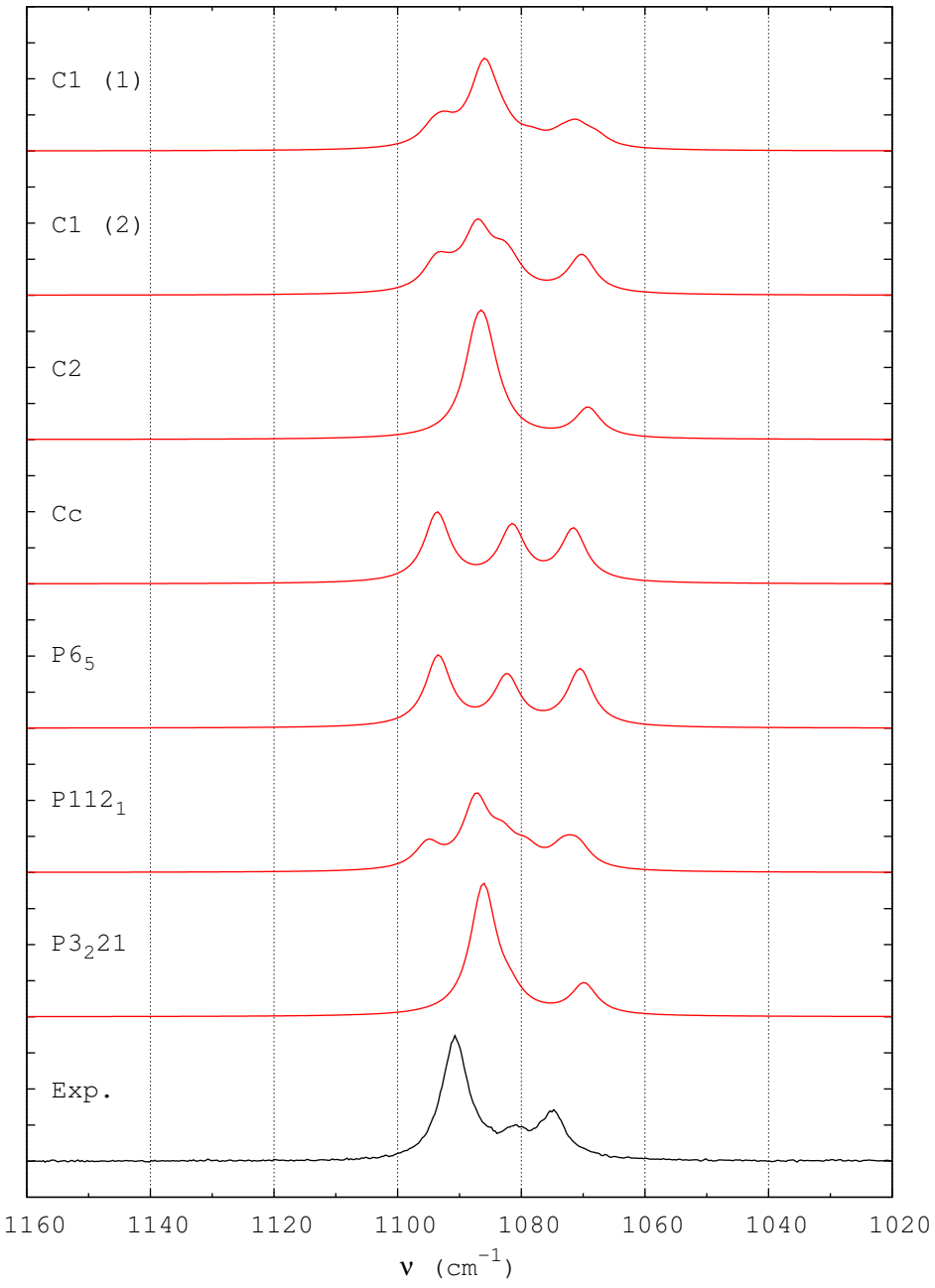

Figure 3 Calculated and experimental Raman spectra of vaterite: expansion of the 900-640 cm^{-1} range (ν_2 and ν_4 modes). The experimental spectrum has been magnified by a factor 2 compared to the computed ones.

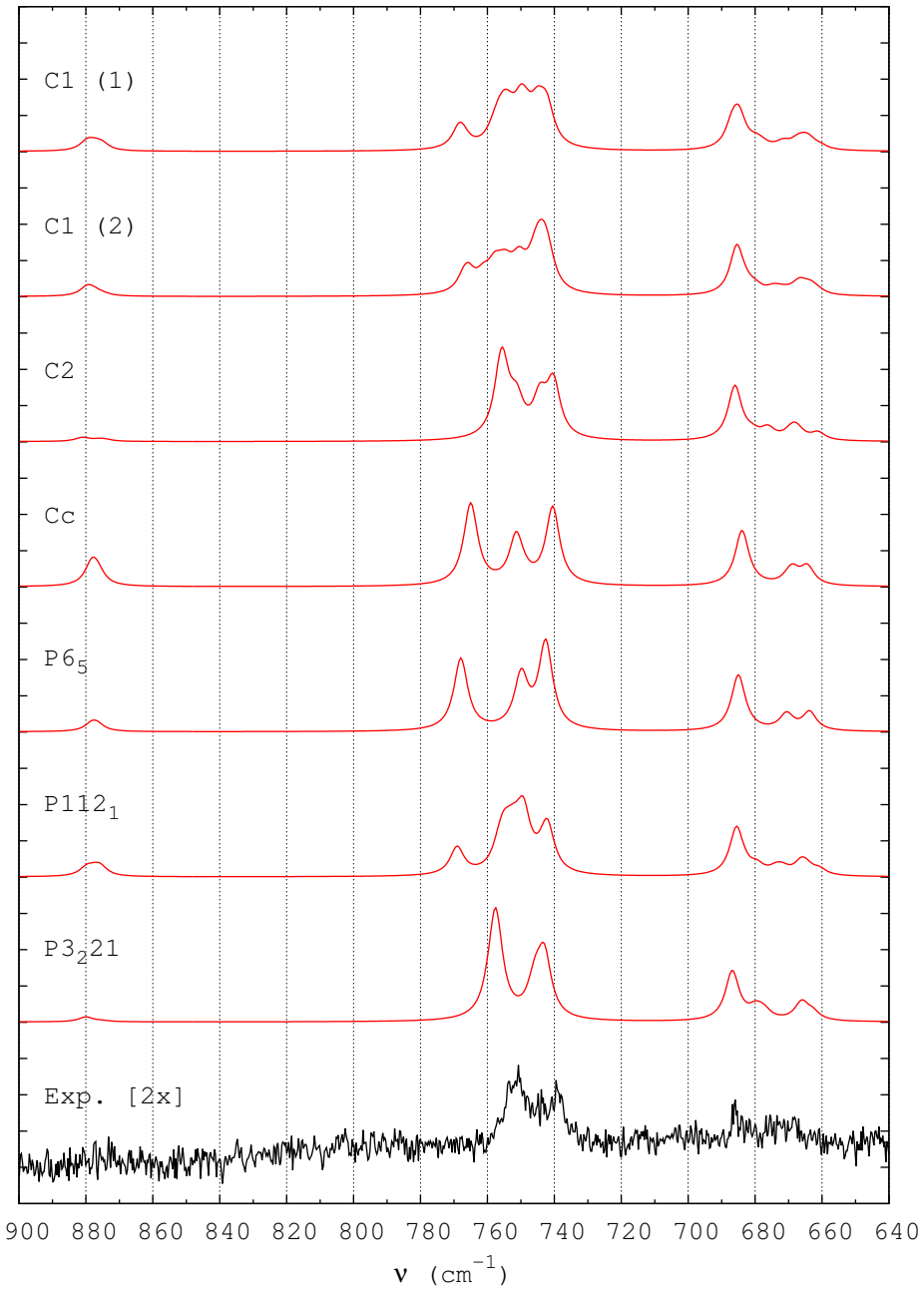

Table of contents

Despite playing a key role in biomineralization, the structure of vaterite is still debated. We show here that only the two most stable models recently proposed as a result of computer simulation, corresponding to two isoenergetic polytypes of vaterite, exhibit Raman spectra that agree with those measured on a collection of geological, biological and synthetic samples.