Structure Determination by Joint Effort of X-ray Powder Diffraction and Quantum Calculations: Crystal Structure and Short Hydrogen Bonding in Pentadecafluorooctanoic Acid Hydrate

Jernej Stare,[1],* Anton Meden,[2],* Dušan Hadži[1],*

[1] Theory Department, National Institute of Chemistry, Ljubljana, Slovenia

[2] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia

* Corresponding authors' e-mail addresses: jernej.stare@ki.si, tone.meden@fkkt.uni-lj.si, dusan.hadzi@ki.si

---

THIS PAPER IS DEDICATED TO DR. BISERKA KOJIĆ-PRODIĆ ON THE OCCASION OF HER 80TH BIRTHDAY

Abstract: A precise X-ray crystal structure determination of systems that exist in a polycrystalline form often poses a challenge due to several factors that limit the resolution of the diffraction measurement. The location of atoms is usually determined at significantly lower precision than the unit cell parameters, rendering the complete characterization of the structure difficult. This is particularly pronounced when precise location of hydrogen atoms is required, for instance in ionizable biomolecules or in hydrogen bonds. In such cases periodic quantum (DFT) calculations may crucially assist structure determination, because they can reliably predict the location of atoms, provided that the unit cell parameters, the space group and tentative atomic positions are known. In this work we present DFT-assisted structure determination of pentadecafluorooctanoic acid hydrate, a benchmark system featuring short hydrogen bonds ($R_{O\cdots O} \approx 2.5$ Å) between acid and water molecules. While Rietveld refinement based solely on powder diffraction data cannot reliably resolve the location of crystal water molecules and even less so the position of hydrogen atoms in the network of hydrogen bonds, periodic DFT optimization yields several minimum energy structures suitable for further refinement. Due to the low experimental resolution and similarity, comparison between model and experimental powder diffraction pattern can barely distinguish between certain structure candidates provided by DFT calculations, but it can rule out those featuring larger misfits. The proposed structure solution is delivered from a tandem application of structure determination from powder diffraction data and DFT optimization, the former providing the unit cell parameters and estimated atomic positions that are finely tuned by DFT. The present strategy can in principle be generalized to other examples of structure determination at relatively low resolution, such as is often the case with biological macromolecules.

Keywords: powder diffraction, structure determination, periodic DFT calculations, hydrogen bonding.

INTRODUCTION

CRYSTAL structure determination from single crystal X-ray diffraction (XRD) data is a well established and powerful method, providing the key data to understanding properties of various materials (structure – properties relationship). There are however, many cases where the suitable single crystals cannot be grown. For polycrystalline materials it is possible to apply the powder diffraction method for structure determination and analysis, but there are strong limitations. The powder pattern is namely the projection of the three-dimensional diffraction image that contains the structural information to one dimension only. This leads inevitably to reflection overlap in the powder pattern and the information content is severely reduced. This becomes even more a problem in cases, where the quality of the crystallites is low and the powder pattern exhibits broad peaks and/or the diffraction becomes diffuse already at low angles (large d-values) where the peaks are hidden in the background from where the information cannot be extracted.
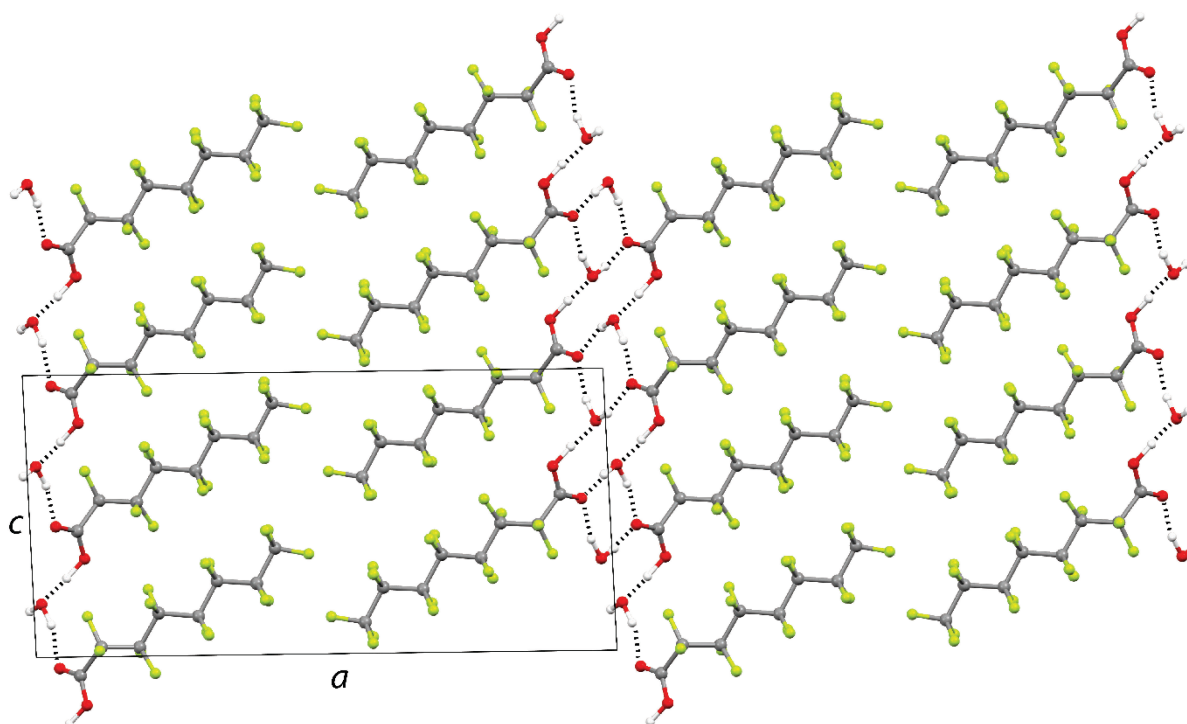
Figure 1. Crystal structure of PDFO·H$_2$O viewed along the *b* axis and with labeled unit cell. H-bonds are denoted with dashed lines.

Limitations of X-ray structure determination are particularly pronounced with hydrogen atoms due to the fact that in the vicinity of hydrogen nucleus the electron density is low. Severely underestimated C–H, N–H and O–H distances in crystalline organic compounds are com-monly reported in many XRD studies. These short-comings can be experimentally avoided by neutron diffraction, which is far more sensitive to hydrogen atoms than XRD, because the neutrons are scattered on nuclei and the scattering power for neutrons is not proportional to the atomic number. In the case of low quality of crystals or crystallites in the polycrystalline samples this advantage is seriously hindered and methods beyond diffraction need to be used (for example spectroscopy or calculations). Examples of systems requiring special care in this regard are ionizable groups in biomolecules, since many of their functional properties are governed by protonation states. Another structural motif requiring precise determination of hydro-gen positions is hydrogen bonding (H-bonding) because of its peculiar (and to a large extent poorly understood) dy-namics, particularly in the case of short and strong H-bonds.

Periodic quantum calculations, mainly based on the Density Functional Theory[1,2] (periodic DFT), offer a valuable support for precise determination of atomic positions that can hardly be resolved by crystallographic techniques at low resolution. Based on the fundamental physical background and free of empirical bias, quantum calculations can reliably predict the structure and other characteristics of molecular systems in various phases, including the crystalline solid state. Although crystal-lography and quantum chemistry are to a large extent complementary, they share a perspective in that at high resolution the experimental electron density is expressed in fine detail, thus approaching the first principle representation defined by quantum mechanics. Therefore, linking XRD and quantum chemical calculations appears like a natural evolution of both methodologies. Indeed, the recently developed quantum-based crystallographic refine-ment (Q|R) uses quantum calculations as inherent part of the structure determination protocol.[3] In addition, periodic DFT calculations have been routinely employed in crystal structure validation, featuring high accuracy and reliability.[4] Our recent research includes several examples of using periodic DFT as supporting tool for structure determination.[5–10]

In this work we scrutinize the crystal structure of pentadecafluorooctanoic acid hydrate (PDFO·H$_2$O). Accor-ding to the structure solution from powder XRD, the system crystallizes in alternating hydrophobic and polar bilayers, constituted by the hydrophobic perfluorinated alkyl chains and the polar carboxylic groups H-bonded to water, respectively (Figure 1). PDFO·H$_2$O is interesting for several

features, in the first place for its short ($R_{O \cdots O} \approx 2.5$ Å) H-bond formed between the acid's carboxylic OH group as donor and water oxygen as acceptor. Together with the longer H-bonds with water as donor and the acid's carbonyl group as acceptor a complex, cross linked network of H-bonds is established, including cyclic, chain, and helical motifs. The H-bonded network of PDFO·H_2O resembles the one of oxalic acid dihydrate, one of the most popular and most thoroughly elucidated H-bonded systems,[11–15] as well as the analogous acetylene and diacetlyene dicarboxylic acid dihydrates.[16,17] The resemblance between PDFO·H_2O and these systems extends to the infrared and Raman spectra that exhibit unique features related to proton dynamics.[18,19] This is also supported by ac conductivity measurements which give evidence for the polaronic character of conductivity.[20,21] Due to its sizable hydrophobic moiety, PDFO is a prototype for lipid membranes.[22] Alongside with other perfluorinated carboxylic acids, it is used as surfactant and emulsifier in the production of fluoropolymers; it is also a persistent organic pollutant.[23]

One of the most interesting features of PDFO·H_2O and related systems is that the shortness of its H-bonds in the crystal is governed by the strong cooperative effect. Namely, as already demonstrated for oxalic acid dihydrate,[24] infrared spectroscopy and cluster calculations suggest that in aqueous solution the H-bonds are much weaker and longer ($R_{O \cdots O} \approx 2.65$ Å). In the crystal, however, the ordered H-bonded motifs cause strong additional polarization which provides substantial enhancement of the H-bonds, resulting in sizable shrinking (by ≈ 0.15 Å). Similar cooperative effect in the crystals has been reported for a number of other systems.[25–28]

Because of its enigmatic proton dynamics and its close relation to a popular benchmark system, detailed structure of PDFO·H_2O is of a great interest in efforts to elucidate important aspects of short H-bonding, mainly in the context of vibrational spectroscopy. However, structure determination by XRD faces severe limitations caused by the fact that single crystals of PDFO·H_2O could not be prepared, and that the crystallographic quality of the powdered sample is low. This limits the effective 2θ range of the measurement to just above 50 degrees, resulting in a resolution of ≈ 1.7 Å, allowing for a coarse placement of the system in the unit cell, but leaving finer details, including the location of crystal water molecules and the entire H-bond network, uncertain. In essence, precision of the present measurement is close to what is common for biological macromolecules studied by single crystal XRD. The scope of this work is to circumvent the aforementioned limitations and determine the crystal structure of PDFO·H_2O by synergistic use of powder XRD and periodic DFT techniques.

METHODS AND MODELS

X-ray Powder Diffraction

Data were collected on a STOE STADI P diffractometer in Debye-Scherrer geometry, equipped with primary germanium monochromator ensuring monochromatic $CuK_{\alpha 1}$ radiation ($\lambda = 1.54056$ Å) and position sensitive detector, covering angular range of 4 $^{\circ}2\theta$. Sample was filled into a 0.3 mm diameter capillary and the powder pattern was collected from 2 to 70 $^{\circ}2\theta$ in steps of 0.02 $^{\circ}$ and a total integration time per step of 3600 seconds. The crystallinity of the sample proved to be rather low and despite the long integration time, ensuring reasonable counting statistics and signal to noise ratio, there were no visible diffraction peaks above 53 $^{\circ}2\theta$ so that the range from 2 to 53 $^{\circ}2\theta$ was finally used for indexing and structure solution.

Peak positions were determined using the PANalytical HighScorePlus software[29] and various programs built into this software were applied for indexing the powder pattern (Treor,[30] Ito,[31] Dicvol[32]). The same monoclinic unit cell with high figure of merit (above 20 in all cases) and no unindexed peaks was found by all three programs. According to systematic absences, the most probable space group was $P2_1/c$. The unit cell volume and expected density of the material corresponded to one PDFO molecule and one molecule of water in the asymmetric unit.

Various computer programs were applied for structure solution trials and finally the most suitable solution was found by Topas Academic using the following strategy. First the profile fit without the structural model was done to determine the unit cell and peak profile parameters. Pseudo-Voigt profile with the asymmetry correction was applied. After achieving a good fit, the unit cell and peak shape parameters were fixed throughout the structure solution process. Then the simulated annealing (SA) process was started with one PDFO molecule and one water molecule (simulated with only one O atom, having the population parameter of 1.25 to account for two electrons from hydrogen atoms) in the asymmetric unit. The PDFO molecule was constructed to have similar bond lengths angles and torsions as in the structures found in the CSD.[33] During SA the position and orientation of the PDFO molecule, the position of water oxygen atom as well as strongly restrained (allowed only to move for ± 10 % around the initial value) bond lengths, angles and torsions within the molecule were varied. Many SA runs with tens of thousands of steps were performed and in a few cases a reasonably looking structure was found, which means that the polar "heads" of the PDFO molecules were packed together, intermolecular distances were almost acceptable, and the position of the water molecule made some

sense. The profile fit, however was not good at all (R_{wp} > 25%). Closer inspection of the misfit revealed that the major differences between calculated and observed pattern are due to h 0 0 reflections, which made us believe that there is a problem of preferred orientation (PO). The crystallites habit (very thin tiny leafs) would allow that but it is rather uncommon that this phenomenon is to such an extent present in the samples packed into a capillary. Despite doubt, refinement of the March-Dollase[34] PO parameter in the direction of 1 0 0 was applied and the R_{wp} decreased to around 15 % (final PO parameter was around 0.75). In the following SA runs the PO parameter was varied along with the structural parameters and the aforementioned structure packing was reproduced more frequently.

It was, however, not possible to proceed from the coarse solution to a decent Rietveld refinement of the structure, because of the limited quality of the pattern (suffering from preferred orientation and extending only to 53 º 2θ, corresponding to the d-value of about 1.73 Å). The coarse structural model was therefore submitted to DFT calculations for improvement.

Computational Methods

The crystal structure of PDFO·H_2O was optimized by periodic DFT calculations using two distinct programs, namely CRYSTAL06[35,36] and VASP v. 5.3.5.[37–41] While the treatment with CRYSTAL is based on localized atomic basis functions, VASP is a plane-wave DFT code. The CRYSTAL calculations were performed at the B3LYP/6-31G(d,p) level of theory.[42–44] VASP calculations utilized the PBE functional[45] corrected for dispersion interactions by the DFT-D3 method of Grimme,[46] a plane-wave basis set with the kinetic energy cutoff of 500 eV, and Projector Augmented Wave (PAW) atomic pseudopotentials.[47] The integrals in the reciprocal space were evaluated on a Monkhorst-Pack k-point mesh[48] of 1 × 4 × 2 points in both CRYSTAL and VASP calculations. All optimizations proceeded by rigorously considering periodicity and following the constraints of the experimental $P2_1/c$ space group. While the majority of optimizations were performed in a fixed unit cell using cell parameters acquired by diffraction, in some cases the cell parameters were also subject to optimization. The above described methodology was validated for the related system of anhydrous PDFO of which the crystal structure was published recently,[49] as well as of oxalic acid dihydrate,[14] yielding very good agreement between the optimized and measured structure. However, in the case of oxalic acid dihydrate VASP exhibits a tendency of overestimating the shortness of the H-bond by about 0.02 Å. In this regard CRYSTAL appears to be more accurate, reproducing the experimental O···O distance to quantitative precision.[18]

At the first stage, optimizations started from the coarse solution described in the previous paragraph, and the crystal water molecule was intuitively placed at several positions and in several different orientations. Different positions of hydrogen atoms were also considered, e.g. keeping the PDFO molecule neutral or moving the proton to the acceptor water molecule. The calculated powder patterns of the resulting optimized structures were then compared against the measured one and evaluated on the basis of their relative energies and structural details.

RESULTS AND DISCUSSION

Profile refinement without structural model converged to a satisfactory fit (R_{wp} = 4.85 %, Figure 2) from which reliable unit cell parameters of a = 22.826(5), b = 5.5204(3), c = 11.115(4) Å and β = 91.47(4) º were determined.

Among the DFT-optimized structures we firstly considered (and ruled out) those in which the unit cell parameters were optimized alongside the atomic positions. In most cases, the unit cell shrinks on optimization by approximately 50 Å³ (≈3–4 %). The mismatch most likely originates from subtle sources of error inherent to the applied numerical model, i.e. DFT and basis set. Although the difference in volume is far from being dramatic – the agreement can eventually be denoted as good – it has a sizable impact on the pertinent model diffraction pattern. When the unit cell changes, the diffraction peaks shift significantly, severely worsening the agreement with the experimental diffraction pattern. One such example is presented in Figure 3. Apparently, sensitivity of the diffraction pattern to the unit cell parameters exceeds the accuracy of DFT optimization of the same parameters by a large margin. In addition, among all the quantities acquired from powder diffraction, unit cell parameters and symmetry are the most reliable and accurate. Therefore we proceeded with models in which the unit cell was fixed to the experimental values.

DFT optimization of the models, constructed by varying the position of the crystal water molecule and the location of hydrogen atoms, resulted in ten energy-minimized structures, of which five were obtained by CRYSTAL and five by VASP. In the following text these structures are denoted as C1-C5 and V1-V5, respectively. Structures with the same number index are in general more similar to one another than to structures with other indexes, hence they can be considered in pairs C1-V1, C2-V2, and so on. Main features of the structures are listed in Table 1 and displayed in Figure 4.

The optimized structures displayed in Figure 4 differ noticeably in the H-bond pattern as well as in conformation of the carboxylic group. The COOH group assumes either 'Z' orientation (i.e. hydrogen is on the same side of the
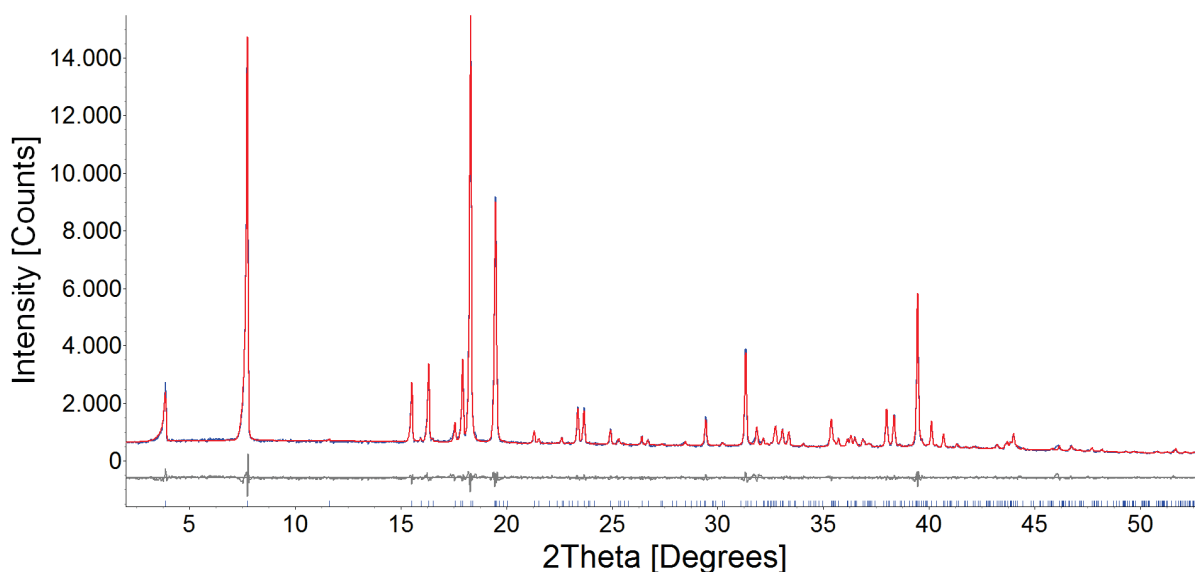
Figure 2. Profile fit of the PDFO·H_2O XRD powder pattern. The measured profile is blue, the calculated one is red, and their difference is the grey line below. Vertical blue bars denote reflection positions.

carboxylic carbon atom as the carbonyl oxygen atom), which is common for carboxylic groups, or the opposite 'E' orientation disfavored in terms of energy[50] and occurring only in special cases such as strong intramolecular H-bonds.[51] For that reason we denote the 'Z' orientation 'common' whereas 'E' is 'uncommon' (Figure 4). Likewise, in the carboxylic group the carbonyl (C=O) oxygen is by far more common H-bond acceptor than the hydroxylic

(C—OH) oxygen, hence the term 'common' and 'uncommon' for those H-bonds, respectively (Figure 4).

In this regard the most commonly looking structure is 1 (C1, V1) in which the OH group of PDFO donates a H-bond to the neighboring water molecule, whereas its C=O group accepts two H-bonds, altogether resembling the structure of oxalic acid dihydrate.[14] Structure 2 (C2, V2) is derived from 1 by migration of the proton from PDFO to
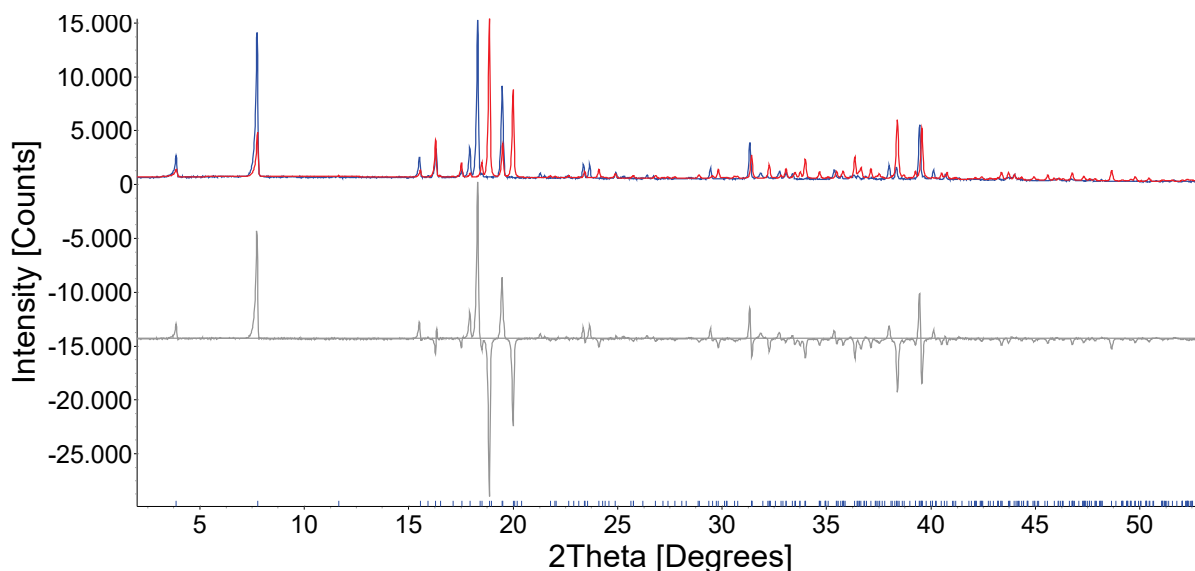

Figure 3. Example of a profile comparison based on a DFT-optimized model with a variable unit cell (structure V1, see text and Figure 4). The measured profile is blue, the calculated one is red, and their difference is the grey line below. Vertical blue bars denote reflection positions. The substantial difference between the measured and the calculated profile apparently originates from the variable unit cell during structure optimization.

Table 1. Relative energies, type of the short hydrogen bond with acid as donor and water as acceptor ('neutral' denotes no proton transfer whereas 'zwitterionic' indicates that the proton is transferred from the donor to the acceptor site), and selected hydrogen bond distances of the ten structures of PDFO·H_2O obtained by geometry optimization by CRYSTAL and VASP. For labeling and geometry of the structures, see text and Figure 4. Note that the energies are given relative to the structure with the lowest energy both (and independently) for CRYSTAL and VASP calculations

index	program	relative energy (kcal/mol)	HB type	R_{OO} (Å)	R_{OH} (Å)
1	CRYSTAL	0.00	neutral	2.496	1.056
	VASP	0.00	neutral	2.451	1.127
2	CRYSTAL	0.44	zwitterionic	2.455	1.385
	VASP	1.26	zwitterionic	2.468	1.389
3	CRYSTAL	4.20	neutral	2.446	1.081
	VASP	0.80	zwitterionic	2.573	1.562
4	CRYSTAL	4.28	neutral	2.479	1.053
	VASP	0.77	zwitterionic	2.613	1.635
5	CRYSTAL	11.72	neutral	2.554	1.018
	VASP	15.24	neutral	2.517	1.049

water, accompanied by a substantial rotation of the COO$^-$ group. On the other hand, structures 3 (C3, V3), 4 (C4, V4)
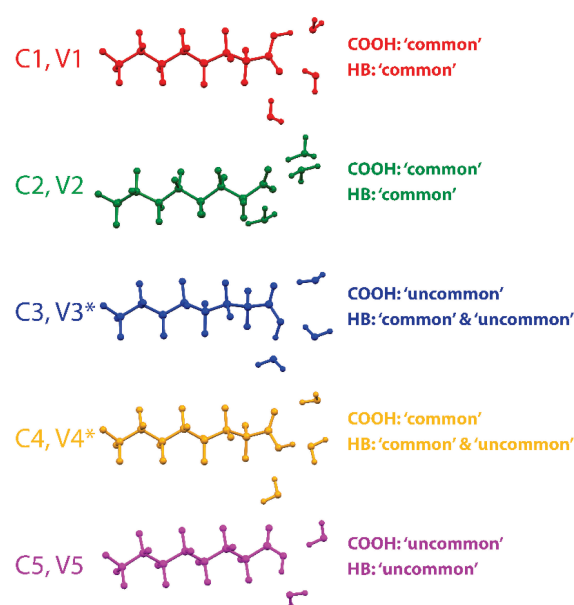

Figure 4. Hydrogen-bonded complexes of PDFO and water molecules, extracted from structures optimized by CRYSTAL and VASP. Together with a PDFO molecule are displayed all water molecules involved in hydrogen bonding with that PDFO molecule. Note that each complex corresponds both to CRYSTAL and VASP optimization, because the alignment of the molecules is very similar regardless of the program. An asterisk (*) indicates that in V3 and V4, unlike the actually displayed geometry, the proton is transferred from PDFO to water molecule. Conformation of the carboxylic group and the hydrogen bond pattern is denoted for each structure as 'common' or 'uncommon'. See text for explanation and Table 1 for selected features.

and 5 (C5, V5) feature either the less common orientation of the carboxylic group, or the rather unusual H-bonding of water to the C—OH oxygen, or both; implying that these structures are less likely to be correct. This is further supported by their relative energies, at least those computed by CRYSTAL. In addition, structure 5 appears to be noticeably less stable due to the fact that a PDFO molecule forms only two H-bonds, whereas it forms three in other structures.

Differences also exist between the related structures optimized by different programs. At a first glance these differences may be less pronounced than those within the H-bonded network (Figure 4), but they may actually be crucial for the agreement with the XRD measurement. By graphically overlaying the structures we found out that the pairs of structures C1/V1, C2/V2 and C5/V5 are very similar to each other with the differences mainly in the hydrophobic part of the PDFO molecule (the CF_3 group), where the position of atoms in the unit cell differ by about 0.1 Å at most. The differences in the pairs C3-V3 and C4-V4 are slightly larger, mainly in the positions of atoms at both ends of the PDFO molecule (CF_3 and COOH groups, including proton transfer in V3 and V4 which does not occur in C3 and C4 – see Table 1), but the differences are still relatively small (most significant ones are about 0.2 Å for the atoms of the PDFO molecule). Worthy to note is that the VASP-optimized O···O distance is by about 0.13 Å longer in V3 and V4 than in the corresponding C3 and C4 structures optimized by CRYSTAL. Differences in H-bond geometry in other cases were significantly smaller. The only pair of models that differed significantly in the position and orientation of the water molecule was C4-V4, where the difference in position of water was about 0.4 Å, while in all other pairs the difference was less than 0.1 Å.

The listed H-bond distances (Table 1) reveal a trend that in comparison to CRYSTAL, VASP calculations predict a
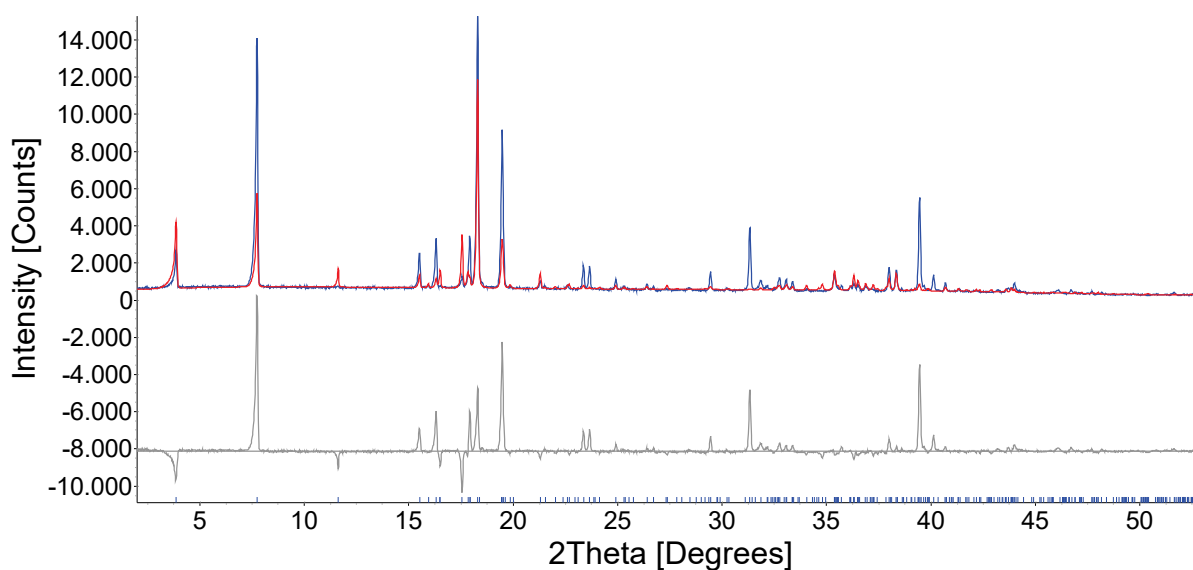
Figure 5. A profile comparison based on a DFT-optimized model C2. The measured profile is blue, the calculated one is red, and their difference is the grey line below. Vertical blue bars denote reflection positions. The substantial difference between the measured and the calculated profile apparently originates from the mismatch of the intensities and is due to wrong packing of the molecules.

larger tendency of the proton to migrate away from the donor oxygen atom. For instance, structures V3 and V4 feature proton transfer ('zwitterionic' form) whereas C3 and C4 remain 'neutral'. In the case of 'neutral' H-bonds (C1/V1, C5/V5) the O—H distance is longer with VASP calculations and, consequently, the O···O distance is shorter (the shortening of the donor···acceptor distance when the proton approaches the midpoint is commonly observed and results from the coupling of the coordinates). We believe this trend is relevant for the proposal of the most favorable structure; as explained below, the choice will be made between C1 and V1.

Structure comparison within the series C1-C5 and V1-V5 reveals that the packing of the PDFO molecules is similar in the cases of C1/V1, C3/V3 and C4/V4. These structures differ in the position of water molecule and hydrogen bond pattern (see above), but the match in packing is apparently crucial for the agreement of these structures with the experimental diffraction data. The packing of PDFO molecules is, however, significantly different in the models C2/V2 and C5/V5. These two models were eliminated from further consideration by comparing their calculated powder patterns with the measured one. The different packing is reflected in substantial differences in the patterns, leading to the conclusion that the models C2/V2 and C5/V5 are not relevant (comparison for C2 is given in Figure 5). The differences between the calculated and measured patterns were even more pronounced in the case of C5/V5. It should

be noted that due to the low resolution of the measurement, the use of the experimental diffraction pattern for evaluation of the computed structures is limited. As a result, comparison with the measured pattern can only evaluate the match of large motifs such as the perfluorinated alkyl chain, whereas finer details such as the H-bond network (let alone the position of hydrogens) can barely be detected, if at all. Therefore, structures C2/V2 that represent the most plausible possibility of proton transfer have not been ruled out due to unfavorable location of the proton but rather due to unfavorable orientation of the hydrophobic tail. An overlay of C1/V1, C2/V2 and the coarse structure model derived from XRD displayed in Figure 6 clearly shows that the C1/V1 structure matches the experimental one much better than C2/V2.

The powder pattern calculated from the C1 model matched best to the measured one (Figure 7). R_{wp} of 12.7 % was achieved only by fitting the scale factor, overall B and March-Dollasse PO parameter, while profile shape and unit cell parameters were kept at the values obtained with profile fitting. Following the same procedure, models C3 and C4 gave slightly worse R_{wp} values (15.1 and 14.5 % respectively).

We tried to improve the model C1 using Rietveld refinement applying strong restrictions to the geometry of the PDFO and water molecules. The Z-matrices of both molecules were constructed so that they exactly resembled the internal structure optimized in the C1 model. Both molecules were then placed at the same position in the
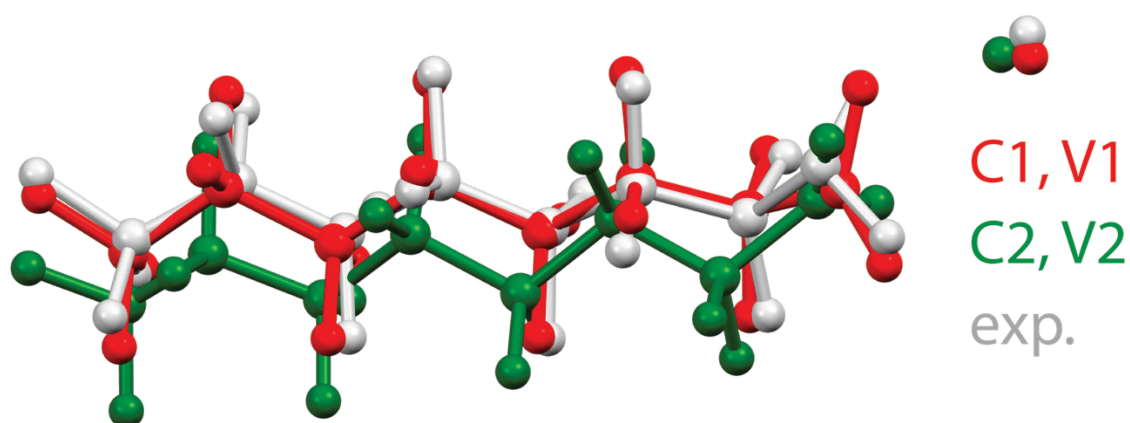
Figure 6. Overlay of the asymmetric unit of selected DFT-optimized structures (for labeling see text and Figure 4) and the coarse structure derived from powder XRD ("exp."). Hydrogen atoms are not shown for the sake of clarity.

same orientation as in C1. From this starting point the position and orientation of both molecules were refined without restraints, while the internal structure was strongly restrained (bond lengths were allowed to vary for ±0.02 Å, bond angles for ±3 ° and torsion angles for ±5 °). The fit improved considerably (R_{wp} dropped from 12.7 to 8.5 %), but subsequent crystallochemical analysis of the refined structure and calculation of its energy revealed that the structure was in fact "spoiled" by the described refinement procedure. Although the molecules were moved only slightly, the H-bonds pattern was destroyed (disregarding the positions of hydrogen atoms) and also the intermolecular F-F contacts in the hydrophobic part of the

structure were less similar to those in the related structures (for example anhydrous PDFO[49]). This was more or less expected due to the limited quality of the powder pattern, as described above. Therefore, we conclude that the best option for the final structural model proposed in this work is a DFT-optimized geometry. For this sake judgement has to be made among C1/V1, C3/V3 and C4/V4 structures. For several reasons the C1/V1 pair appears to be the one of our choice. First, it has the lowest energy both by CRYSTAL and VASP. For the other two structures the agreement between the programs is not that good, i.e. CRYSTAL predicts noticeably larger energy difference between structures 1 and 3 or 4 than VASP (Table 1). Second, for C1/V1 the
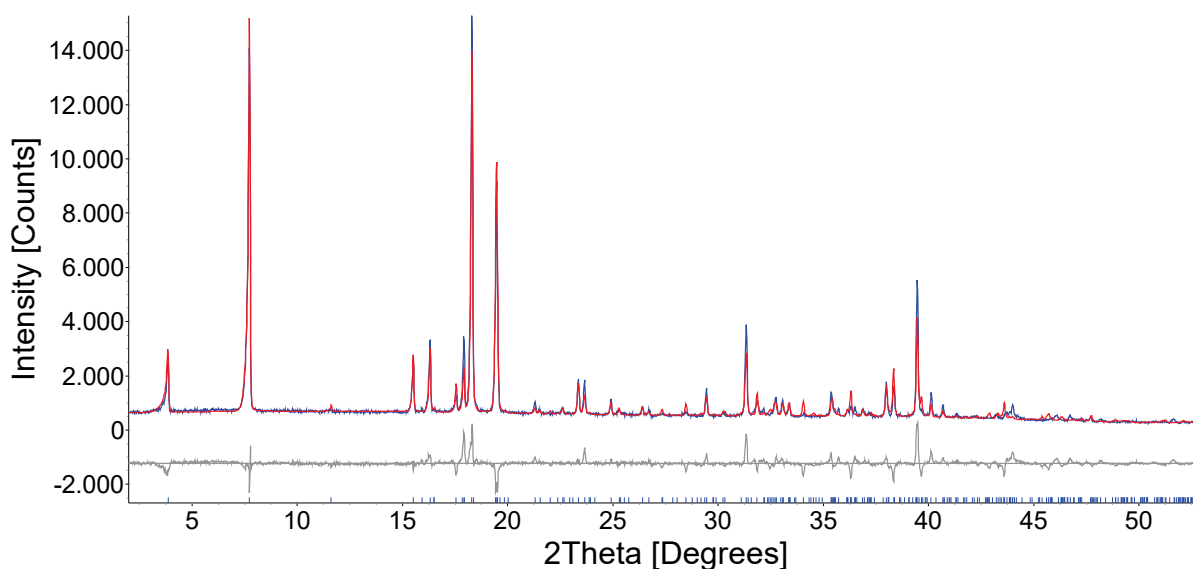

Figure 7. A profile comparison based on a DFT-optimized model C1. The measured profile is blue, the calculated one is red, and their difference is the grey line in the middle. Vertical blue bars denote reflection positions. The difference between the profiles is small enough to confirm correct packing.
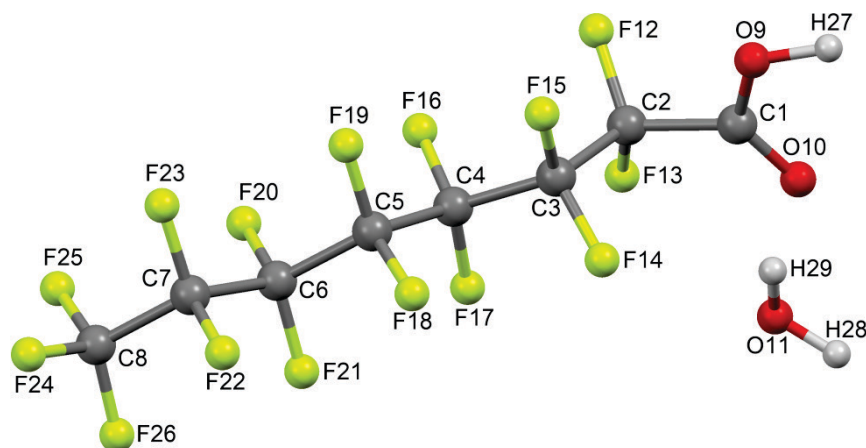
Figure 8. Asymmetric unit and labelling scheme for PDFO · H$_2$O.

agreement between CRYSTAL and VASP extends to the structure whereas structures 3 and 4 are noticeably different between the programs, particularly in the H-bonded moiety. Third, the C1/V1 structure looks by far most 'common' in terms of geometry of the COOH group and the H-bonding network (Figure 4). On its own this is a weak argument, but it is supported by the fact that the spectra of PDFO·H$_2$O remarkably resemble those of oxalic acid dihydrate in peculiar details related to the H-bond, hence the assumption that these systems have similar structure is reasonable. Resemblance is granted with the C1/V1 structure, but not at all with C3/V3 or C4/V4. Finally, although the three pairs of candidates exhibit similar match of their diffraction patterns to the experimental one, the agreement of C1/V1 is slightly better than of the other two pairs.

Despite the C1 and V1 models are about equivalent in many aspects, as discussed above, we tend to believe that C1 represents a better approximation of the real structure than V1. Although the difference of 0.045 Å in the O···O distance between C1 and V1 is probably insignificant for the agreement with the experimental diffraction pattern, the shorter O···O distance optimized by VASP (R_{OO} ≈ 2.45 Å) puts the H-bond into a whole different class, bringing it close to the shortest H-bonds known of this type.[10,51–54] However, the spectral resemblance with oxalic acid dihydrate (R_{OO} ≈ 2.49 Å) does not support this. Importantly, while CRYSTAL optimization at the same level of theory reproduces the H-bond geometry of oxalic acid dihydrate very accurately,[18] VASP overestimates its shortness by about 0.02 Å. Thus, by analogy, the C1 structure is probably more accurate than V1, and the suggested O···O distance is almost identical to that of crystalline oxalic acid dihydrate.[14]

The proposed crystal structure pf PDFO·H$_2$O consists of double layers of PDFO molecules, where water molecules are inserted into the hydrophillic part, as depicted in Figure 1. The asymmetric unit with the atom labelling scheme is presented in Figure 8 and the atomic coordinates, obtained by energy minimization (model C1) are given in Table 2.

Molecular geometry of the PDFO molecule in the proposed structure is within expectations. The C=O (carbonyl) bond length is 1.23 and C–O (hydroxyl) 1.289 Å, which is well comparable to anhydrous PDFO,[49] with 4 molecules in the asymmetric unit, where the corresponding bond lengths range from 1.228 to 1.241 and from 1.288 to 1.331, respectively. The C–C distances in the PDFO chain are all very close to 1.56 Å (from 1.557 to 1.565 Å), which is in average a little longer than the average in the anhydrous PDFO (1.54 Å), but well within the range found therein (1.510 to 1.576 Å). Similar is true for the C–F distances where we observe a small difference between the terminal –CF$_3$ group and –(CF$_2$)– groups along the chain. In our proposed model they are between 1.337 and 1.342 Å for the former and between 1.352 and 1.358 Å for the latter, while the corresponding averages and ranges for the anhydrous PDFO are 1.33 Å (1.31 to 1.35 Å) and 1.35 Å (1.31 to 1.37 Å), respectively.

Comparison of the bond angles also reveals close similarity between the PDFO molecule in our proposed model and those in the structure of anhydrous PDFO. The C1–C2–C3 angle in our model is 112.9 $^\circ$, while the C–C–C angles along the chain are a bit larger – from 113.6 to 114.7 $^\circ$. The corresponding average values and ranges for anhydrous PDFO are 111.0 $^\circ$ (from 110.6 to 111.7 $^\circ$) and 115.4 $^\circ$ (from 112.2 to 118.2 $^\circ$). The F–C–F angles in our proposed model range from 108.6 to 109.3 $^\circ$, which corresponds

Table 2. Coordinates of the atoms in the asymmetric unit of PDFO · H$_2$O. Note that the coordinates are calculated by energy minimization and the standard deviations and atomic displacement parameters are not given

Label	x	y	z
C1	0.084	0.943	−0.061
C2	0.123	1.050	0.044
C3	0.188	0.965	0.040
C4	0.223	1.005	0.162
C5	0.291	0.987	0.150
C6	0.325	0.976	0.274
C7	0.393	0.989	0.266
C8	0.426	0.976	0.390
O9	0.094	1.045	−0.162
O10	0.048	0.783	−0.040
O11	0.019	0.561	0.181
F12	0.121	1.295	0.038
F13	0.100	0.979	0.150
F14	0.187	0.725	0.015
F15	0.214	1.086	−0.049
F16	0.210	1.228	0.204
F17	0.205	0.835	0.241
F18	0.303	0.785	0.087
F19	0.309	1.184	0.089
F20	0.307	1.165	0.342
F21	0.311	0.766	0.330
F22	0.411	0.800	0.199
F23	0.408	1.199	0.213
F24	0.484	0.983	0.369
F25	0.412	1.165	0.459
F26	0.414	0.771	0.448
H27	0.063	1.001	−0.230
H28	−0.003	0.420	0.150
H29	0.034	0.641	0.110

again very well to the anhydrous PDFO, where these angles average close to 109 ° and the range is expectedly broader (from 107 to 111 °).

The torsion angles in the C1–C8 chain in our proposed model are such that the molecule is gradually twisted along the chain in the same direction going from one –(CF2)– group to another, so that the torsion C2–C3–C7–C8 accumulates to 40 °. This is, however, significantly smaller than in the anhydrous PDFO, in which the corresponding average is 72 ° (from 68 to 74 °). We think

that the difference is real, as the torsions do not only depend on the intramolecular interactions, but also on the intermolecular ones, which are different in the two cases (the differences originate in different packing).

The packing of the hydrophobic perfluorinated chains of the PDFO molecules in our proposed model fits well within the expectations (Figure 1). The molecules are aligned together along the chains so that the space is populated densely. Fluorine atoms from one chain fit into the gap between the fluorine atoms of the neighbor chain so that the intermolecular F···F distances are close to the sum of the Van der Waals radii, which is around 3.0 Å. Some contacts are a bit shorter – down to 2.82 Å, in the anhydrous PDFO, the shortest intermolecular F···F distance is 2.72 Å. Similar holds for the packing of the –CF$_3$ groups at the ends of the PDFO molecules.

As it was mentioned, the spectroscopic data suggest a close similarity of the hydrogen bod properties between PDFO·H$_2$O and oxalic acid dihydrate. The H-bond patterns in both structures are presented in Figure 9. The similarity is evident. In both cases the H-bond pattern is centrosymmetric where two water molecules and two carbonyl oxygen atoms form a planar rhombical arrangement with the water hydrogen atoms lying nearly in the plane and forming two donor H-bonds of moderate strength to carbonyl oxygen atoms. In both structures water molecules occupy more distant corners and serve as acceptor for the short H-bond with the acid hydroxyl group. The angle between this bond and the rhombus plane is similar in both cases and the same holds for the angles between the –COOH group planes and the rhombus planes.

The two carboxylic C atoms lie nearly in the plane of the rhombus (in PDFO·H$_2$O the C=O bond is inclined out of the rhombus plane for 8 and in the oxalic acid dihydrate for 5 °. Also the planes defined by the two –COOH groups sitting at each end of the shorter diagonal of the rhombus are close to be parallel to the plane of the rhombus (in PDFO·H$_2$O the deviation is 17 and in the oxalic acid dihydrate 9 °). The planes of the other two –COOH groups (those donating the hydrogen to the shortest H-bond) are, on the contrary, close to being perpendicular to the rhombus plane (in PDFO·H$_2$O the angle between these two planes is 88 and in the oxalic acid dihydrate 75 °).

Periodic repetition of this H-bond pattern in the *bc* plane can be seen in Figure 1. Water molecules serve as linkers between the polar heads of the PDFO molecules within the layer and across the layer. In contrast to similarities presented above, periodicity of the H-bond pattern in oxalic acid dihydrate is different from that in the PDFO·H$_2$O due to different shape, size and number of the –COOH groups and water molecules; however, a more detailed comparison is beyond the scope of this paper.
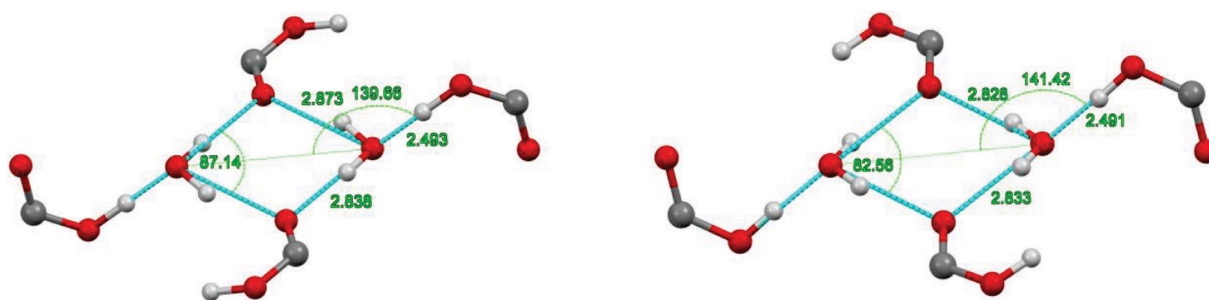
Figure 9. H-bond pattern in PDFO·H$_2$O (left) and oxalic acid dihydrate (right). The distances are given in Å and angles in °.

CONCLUSIONS

The crystal structure of pentadecafluorooctanoic acid hydrate poses a challenge for accurate structure determination because the substance can only be prepared in polycrystalline form of limited quality. Consequently, the resolution of powder XRD is low and only allows for reliable determination of unit cell parameters and space group symmetry. Conformation and packing of PDFO molecules can be deduced at a coarse level, but the hydrogen bond pattern remains beyond the reach of diffraction data alone. In order to overcome this limitation we engaged periodic DFT calculations, which, aided with the experimental structure data, yield several energy-minimized structures differing mainly in the hydrogen bonding network, but also in the packing of the molecules. Comparison between the pertinent model powder diffraction patterns with the experimental one filters the DFT-optimized structure candidates to three. These were further evaluated on the basis of their energy, the commonness of their structural motifs, and on our preliminary findings from vibrational spectroscopy, in the first place the presumed structural analogy with oxalic acid dihydrate. One of the structures matched these criteria significantly better than the other two, therefore it has been proposed as final structure solution. We believe that the presently used methodology consisting of energy minimization using the fixed unit cell and symmetry provided by X-ray powder diffraction data is capable to derive relevant crystal structure models suitable for studying rather fine structural details such as H-bond patterns.

Among the findings worth to be stressed is that sensitivity of the diffraction pattern to the unit cell parameters appears to be noticeably higher than the precision/accuracy at which the unit cell parameters can be optimized by DFT methods. Therefore it is advisable to use unit cells fixed to the experimentally determined values, which are readily available and accurate even with powder samples of poor quality.

Despite the present study shows that structures with proton transfer are in worse agreement with the powder diffraction measurement at given conditions, we need to note that our recent vibrational spectroscopy studies give evidence that proton transfer may occur at temperatures below 150 K, the process probably being driven by polarons.[55] In this sense, the C2/V2 structure may also represent plausible structure solution. This requires clarification and remains a challenge for future work, including variable-temperature structure determination by XRD. In addition, the proton dynamics and the possible tautomeric equilibrium could be elucidated by molecular dynamics simulation based on first-principle forces derived from DFT treatment ("*ab initio* MD").

The present study confirms the predictive power of quantum chemistry methods and their applicability in crystallographic studies, which can possibly lead to the expansion of research combining crystallographic and theoretical methods in the future. One of the possible challenges is improved structure of crystalline biomolecules that often share the problem of limited resolution.

REFERENCES

[1]   P. Hohenberg, W. Kohn, *Phys. Rev. B* **1964**, *136*, 864.
[2]   W. Kohn, L. J. Sham, *Phys. Rev. A* **1965**, *140*, 1133.
[3]   M. Zheng, J. R. Reimers, M. P. Waller, P. V. Afonine, *Acta Cryst. D* **2017**, *73*, 45.
[4]   J. van de Streek, M. A. Neumann, *Acta Cryst. B* **2014**, *70*, 1020.
[5]   A. Kokalj, R. Dominko, G. Mali, A. Meden, M. Gaberscek, J. Jamnik, *Chem. Mater.* **2007**, *19*, 3633.
[6]   I. Djerdj, J. Popovic, J. Stare, G. Ambrozic, S. D. Skapin, B. Kozlevcar, D. Pajic, Z. Jaglicic, Z. C. Orel, *J. Mater. Chem.* **2012**, *22*, 10255.

DOI: 10.5562/cca3348

Croat. Chem. Acta **2018**, *91*(*2*), 209–220

[7]   E. V. Boldyreva, S. G. Arkhipov, T. N. Drebushchak, V. A. Drebushchak, E. A. Losev, A. A. Matvienko, V. S. Minkov, D. A. Rychkov, Y. V. Seryotkin, J. Stare, B. A. Zakharov, *Chem. Eur. J.* **2015**, *21*, 15395.

[8]   D. A. Rychkov, J. Stare, E. V. Boldyreva, *Phys. Chem. Chem. Phys.* **2017**, *19*, 6671.

[9]   N. Biliskov, B. Kojic-Prodic, G. Mali, K. Molcanov, J. Stare, *J. Phys. Chem. A* **2011**, *115*, 3154.

[10] K. Molcanov, J. Stare, M. V. Vener, B. Kojic-Prodic, G. Mali, J. Grdadolnik, V. Mohacek-Grosev, *Phys. Chem. Chem. Phys.* **2014**, *16*, 998.

[11] A. Lehmann, P. Luger, C. W. Lehmann, R. M. Ibberson, *Acta Cryst. B* **1994**, *50*, 344.

[12] J. M. Robertson, I. Woodward, *J. Chem. Soc.* **1936**, 1817.

[13] E. D. Stevens, P. Coppens, *Acta Cryst. B* **1980**, *36*, 1864.

[14] A. Martin, A. A. Pinkerton, *Acta Cryst. B* **1998**, *54*, 471.

[15] N. Casati, P. Macchi, A. Sironi, *Chem. Commun.* **2009**, 2679.

[16] J. D. Dunitz, J. M. Robertson, *J. Chem. Soc.* **1947**, 148.

[17] J. D. Dunitz, J. M. Robertson, *J. Chem. Soc.* **1947**, 1145.

[18] V. Mohacek-Grosev, J. Grdadolnik, J. Stare, D. Hadzi, *J. Raman Spectrosc.* **2009**, *40*, 1605.

[19] V. Mohacek-Grosev, J. Grdadolnik, D. Hadzi, *J. Phys. Chem. A* **2016**, *120*, 2789.

[20] A. Levstik, C. Filipic, V. Bobnar, I. Levstik, D. Hadzi, *Phys. Rev. B* **2006**, 74.

[21] C. Filipic, A. Levstik, D. Hadzi, *J. Appl. Phys.* **2013**, 113.

[22] D. Matyszewska, K. Tappura, G. Oradd, R. Bilewicz, *J. Phys. Chem. B* **2007**, *111*, 9908.

[23] Y. Fujii, K. H. Harada, A. Koizumi, *Chemosphere* **2013**, *93*, 538.

[24] J. Stare, D. Hadzi, *J. Chem. Theor. Comput.* **2014**, *10*, 1817.

[25] N. Kobko, L. Paraskevas, E. del Rio, J. J. Dannenberg, *J. Am. Chem. Soc.* **2001**, *123*, 4348.

[26] P. Vishweshwar, A. Nangia, V. M. Lynch, *Chem. Commun.* **2001**, 179.

[27] B. F. King, F. Weinhold, *J. Chem. Phys.* **1995**, *103*, 333.

[28] M. D. Esrafili, P. Fatehi, M. Solimannejad, *Comput. Theor. Chem.* **2013**, *1022*, 115.

[29] HighScore Plus, V4.1, PANalytical B.V., Almelo, The Netherlands, **2014**.

[30] P. E. Werner, L. Eriksson, M. Westdahl, *J. Appl. Crystallogr.* **1985**, *18*, 367.

[31] J. W. Visser, *J. Appl. Crystallogr.* **1969**, *2*, 89.

[32] A. Boultif, D. Louer, *J. Appl. Crystallogr.* **2004**, *37*, 724.

[33] C. R. Groom, I. J. Bruno, M. P. Lightfoot, S. C. Ward, *Acta Cryst. B* **2016**, *72*, 171.

[34] W. A. Dollase, *J. Appl. Crystallogr.* **1986**, *19*, 267.

[35] R. Dovesi, V. R. Saunders, C. Roetti, R. Z.-W. Orlando, C. M., F. Pascale, B. Civalieri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, M. Llunell, CRYSTAL06 User's Manual, University of Torino, Torino, **2006**.

[36] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V. R. Saunders, C. M. Zicovich-Wilson, *Z. Kristallogr.* **2005**, *220*, 571.

[37] G. Kresse, J. Hafner, *Phys. Rev. B* **1993**, *47*, 558.

[38] G. Kresse, J. Hafner, *Phys. Rev. B* **1994**, *49*, 14251.

[39] G. Kresse, J. Furthmuller, *Phys. Rev. B* **1996**, *54*, 11169.

[40] G. Kresse, J. Furthmuller, *Comp. Mater. Sci.* **1996**, *6*, 15.

[41] G. Kresse, J. Hafner, *J. Phys. Condens. Mat.* **1994**, *6*, 8245.

[42] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.

[43] C. T. Lee, W. T. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.

[44] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724.

[45] J. P. Perdew, K. Burke, M. Ernzerhof, *Acs. Sym. Ser.* **1996**, *629*, 453.

[46] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, 132.

[47] G. Kresse, D. Joubert, *Phys. Rev. B* **1999**, *59*, 1758.

[48] H. J. Monkhorst, J. D. Pack, *Phys. Rev. B* **1976**, *13*, 5188.

[49] H. Omorodion, M. Palenzuela, M. Ruether, B. Twamley, J. A. Platts, R. J. Baker, *New J. Chem.* **2018**.

[50] M. Pettersson, J. Lundell, L. Khriachtchev, M. Rasanen, *J. Am. Chem. Soc.* **1997**, *119*, 11715.

[51] T. Steiner, A. M. M. Schreurs, M. Lutz, J. Kroon, *Acta Cryst. C* **2000**, *56*, 577.

[52] L. Golic, I. Leban, *Croat. Chem. Acta* **1982**, *55*, 41.

[53] W. Joswig, H. Fuess, G. Ferraris, *Acta Cryst. B* **1982**, *38*, 2798.

[54] K. D. Eichhorn, *Z. Kristallogr.* **1991**, *195*, 205.

[55] V. Mohaček-Grošev, J. Grdadolnik, J. Eckert, J. Stare, D. Hadži, in preparation, **2018**.