J|A|C|S COMMUNICATIONS

Published on Web 02/05/2010

Powder Crystallography by Combined Crystal Structure Prediction and High-Resolution ^{1}H Solid-State NMR Spectroscopy

Elodie Salager,[†] Graeme M. Day,[‡] Robin S. Stein,[†,⊥] Chris J. Pickard,[§] Bénédicte Elena,[†] and Lyndon Emsley[†,*]

Université de Lyon, Centre de RMN à très hauts champs, CNRS/ENS Lyon/UCBL, 5 rue de la Doua, 69100 Villeurbanne, France, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K., and Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, U.K.

Received November 6, 2009; E-mail: lyndon.emsley@ens-lyon.fr

Recent progress in molecular modeling strategies has allowed the development of efficient crystal structure prediction (CSP) methods for molecular crystals based only on the atom connectivities in a molecule.[1] They have reached the point where structures corresponding to the naturally occurring polymorphs are almost always among the predicted structures. However, there are typically 10−100 structures within about 10 kJ·mol^{-1} of the lowest-energy structure, and it can be difficult to correlate these predicted structures with a given powder sample. Here, we show how high-resolution ^{1}H solid-state NMR can provide a reliable experimental determination of a powder crystal structure from a set of *ab initio* predicted candidates, through the example of thymol.

Comparison between calculated and experimental powder X-ray diffraction (PXRD) data are most often used to assess the quality of predicted structures,[2] while potential predicted candidates have been assessed using calculated elastic properties and vapor growth morphology.[3] Alternatively, structure predictions can be used as additional information when PXRD alone does not provide sufficient data to resolve a structure.[4]

Solid-state NMR chemical shifts (CS) have long been known to be sensitive to the structure of molecular crystals.[5] Chemical shifts are increasingly used as a crystallographic tool, possibly in conjunction with other experimental or computational techniques, for organic[6] as well as inorganic[7] samples. Facelli and Grant were the first to combine chemical shift tensor calculations and experimental measurements of ^{13}C to get structural information from a single crystal.[8] Recently, Harper and Grant demonstrated for a carbohydrate that selection of the observed structure from an ensemble of predicted structures was possible using the full ^{13}C chemical shift tensor.[9] Spiess and co-workers have pioneered the use of calculated isotropic ^{1}H chemical shifts to assign experimental spectra and validate structural hypotheses.[10]

In the past five years, solid-state NMR proton−proton spin-diffusion (PSD) experiments have also been shown to be sensitive enough to structure to be combined with molecular modeling in the first step of a complete method of NMR crystallography of powders.[11−13]

Here we present a new approach, based on the efficiency of crystal structure prediction methods and the sensitivity of experimental ^{1}H solid-state NMR to structure. An ensemble of potential structures is predicted *ab initio* (based only on the atom connectivities) for powdered thymol (**1**), a small organic local antiseptic compound chosen to be representative of molecular crystals. The structure of the powder under study is then determined from the ensemble of predicted potential polymorphs by comparison of the experimental proton solid-state NMR data (either isotropic chemical shifts or PSD data) with the same property calculated for each of the potential structures. We show that comparison, without the need for experimental assignment, of calculated and experimental solid-state ^{1}H isotropic chemical shifts (measured at natural isotopic abundance), is sufficient to identify the structure of the powder from among the potential structures. The same selection approach using NMR proton−proton spin-diffusion data is also independently successful in identifying the correct structure.

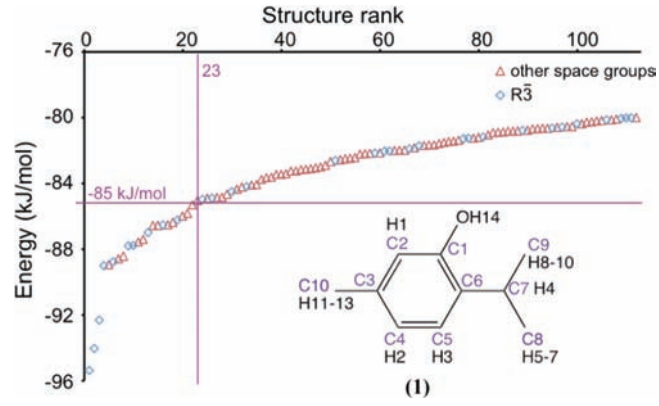

Figure 1. Energies calculated for the 112 lowest-energy predicted crystal structures for thymol (**1**) and the cutoff energy (85 kJ·mol^{-1}) used to determine the ensemble of 23 most reasonable structures.

Starting from the chemical formula of thymol and without any structural hypothesis, crystal structures were predicted *ab initio* using the three-step CSP protocol described earlier for the study of polymorphism in phenobarbital.[14] Trial crystal structures were generated in the most common space groups using the Monte Carlo simulated annealing method and conformations determined from a prior conformational energy scan. These structures were further optimized (unit cell, molecular positions, and conformations) using a molecular mechanics description of inter- and intramolecular forces. The final energy of the 10000 lowest-energy structures is calculated as a combination of a DFT calculation for the intramolecular contribution and an atom−atom model of intermolecular interactions, including an atomic multipole description of the electrostatics. More details of the crystal structure prediction are given in SI. To remove physically unrealistic structures, only structures within 10 kJ·mol^{-1} of the lowest-energy structure were retained (Figure 1). This defined a group of 23 potential polymorphs for thymol. Ten have the same space group as the known, observed

[†] Université de Lyon.
[‡] University of Cambridge.
[§] University College London.
[⊥] Current address: Bruker UK Ltd, Banner Lane, Coventry CV4 9GH, U.K.

10.1021/ja909449k

structure[15] ($R\bar{3}$); the other 13 belong to three other space groups: $P2_1/c$, $P2_12_12_1$, and $C2/c$.

All of these low-energy structures show O−H···O intermolecular hydrogen bonding, with these H-bonds arranged in one of three H-bonding motifs. In $R\bar{3}$, eight structures form cyclic H-bonded hexamers, as in the observed structure; the other two form cyclic trimers. All of the low-energy structures in other space groups contain infinite H-bond chains.

Two categories of experimental ^{1}H NMR spectra from a sample of powdered thymol (see SI for details of the data acquisition) were then used to determine the structure from among the ensemble of predicted candidates: chemical shift measurements and two-dimensional proton−proton spin-diffusion data.

First, chemical shieldings were calculated for the predicted structures without any further geometry optimization using CASTEP,[16] a DFT-based program that takes into account the periodicity of the crystal. Great progress has been made recently in chemical shift prediction, and the GIPAW[17] approach combined with ultrasoft pseudopotentials[18] has been shown to be very accurate for solids.[19] Chemical shieldings σ_{calc} were converted to chemical shifts δ_{calc} using the property $\delta_{calc} = \sigma_{ref} - \sigma_{calc}$. This conversion is done using a reference value, σ_{ref}, to best align the calculated and experimental spectra. In this study, the reference value was determined by a fit between calculated shieldings and experimental chemical shifts for each structure. The average value of σ_{ref} is 31.18 ppm ± 0.28 ppm for proton chemical shifts and 170.1 ppm ± 1.0 ppm for carbon chemical shifts.

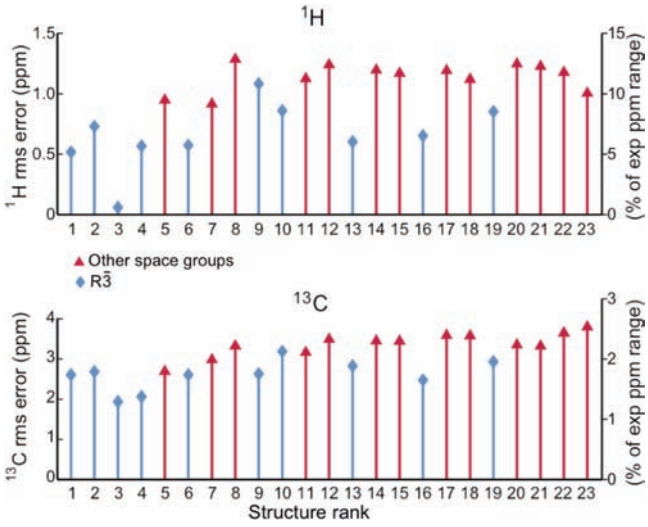

Figure 2. Comparison between DFT calculated and experimental chemical shifts for the 23 structures.

Figure 2 shows the root-mean-square (rms) error between calculated and experimental chemical shifts for both ^{1}H and ^{13}C. Importantly, agreement between experimental and calculated chemical shifts is not correlated with predicted energy (structures are ordered by ascending predicted energy). Thus, experimental isotropic chemical shifts contain information complementary to that contained in the energy models used in the CSP protocol. These data identify the third-lowest energy structure as the one in best agreement with the measured chemical shifts, with an rms error between calculated and experimental chemical shifts of only 0.06 ppm for ^{1}H and 2.15 ppm for ^{13}C.

To estimate uncertainty in the calculated values, we carried out chemical shift calculations for 15 organic compounds (X-ray structures with CASTEP-optimized ^{1}H positions) and found an average rms error of 0.33 ppm (±0.16 ppm) from the experimental values for ^{1}H and 1.9 ppm (±0.4 ppm) for ^{13}C. Differences smaller than these average values can thus be considered insignificant here. It is possible that a slightly different scatter may be obtained for the predicted structures.

It is of particular interest that the second best candidate (structure 1), which is the lowest-energy predicted structure, can be discarded due to a much poorer agreement for ^{1}H (the difference in the rms errors between these structures is higher than 0.33 ppm, estimated uncertainty on ^{1}H). It is noteworthy that the eight structures with the lowest rms error in ^{1}H chemical shifts all have the hexamer H-bonding pattern. Although ^{13}C chemical shifts indicate the third-lowest energy structure as the structure in best agreement with measured ^{13}C chemical shifts, the scatter of the rms errors is within the uncertainty (1.9 ppm), preventing us from drawing a conclusion on the basis of the ^{13}C data only.

Figure 3 shows a comparison of the ^{1}H and ^{13}C experimental spectra and the spectra calculated for the best predicted structure. As an added benefit, assignment of the experimental spectra can be obtained from the calculated spectrum of the correct structure identified from the ensemble of potential polymorphs. In this case, this assignment fits the previously published experimental assignment,[13] except for C8 and C7 which are only 0.6 ppm apart in the experimental spectrum (−0.8 ppm in the calculated spectrum).

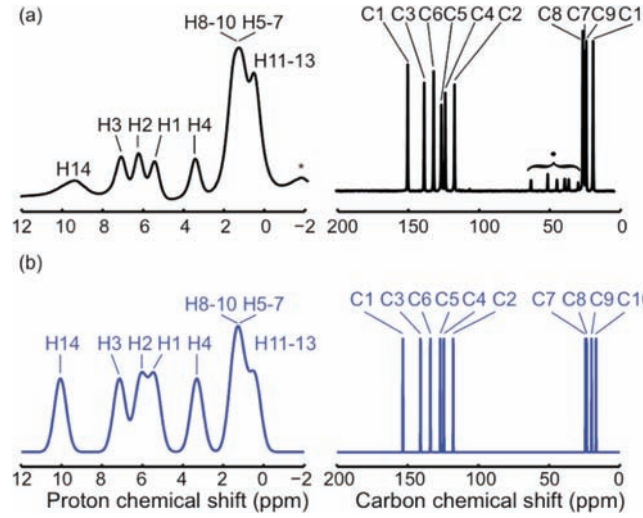

Figure 3. (a) Experimental spectra and (b) spectra calculated for the structure determined from the ensemble of potential polymorphs. The calculated spectra were drawn using a line width of 0.2 ppm for the proton spectrum and 0.1 ppm for the carbon spectrum.

Proton−proton spin-diffusion (PSD) experiments were also tested for their ability to determine the correct structure out of the set of predicted polymorph structures. Experimentally measured PSD build-up curves were compared to curves calculated using a home-written C++ routine, for each of the predicted structures, from the atomic coordinates of the protons included in a 15 Å sphere in the crystal structure around the observed molecule. The principle of the rate matrix analysis and the parameters used to simulate these curves have been published elsewhere.[11] The agreement between calculated and experimental build-up curves (goodness-of-fit coefficient χ^2) is then used to determine which of the predicted crystal structures best describes the experimental data. Figure 1SI, SI, shows the χ^2 coefficients together with calculated energies for the 23 potential polymorphs. Again, NMR data and predicted energy are not correlated, and

in agreement with the chemical shift selection method, the third-lowest energy structure clearly shows the best agreement to the experimental data. The difference in the goodness-of-fit coefficients of the best and the second best structure is higher than the uncertainty and confirms that the third-lowest energy structure is the only one in agreement with the experimental data.

Figure 4 illustrates the structure selected from the ensemble by both methods, as compared with the single crystal X-ray determined structure. The all-atom root-mean-square distance between the two closest molecules in the best predicted structure and the single-crystal X-ray structure[15] is 0.29 Å, illustrating the remarkable accuracy of structure prediction in this case. This can be compared to 0.32 Å for the results of an extensive molecular modeling fit to the PSD data.[13] Note that the result here does not identify the lowest predicted energy structure as the naturally occurring one. The natural structure is the third-lowest in the set of predictions, 3 kJ·mol^{-1} higher in predicted energy than the most stable predicted structure.

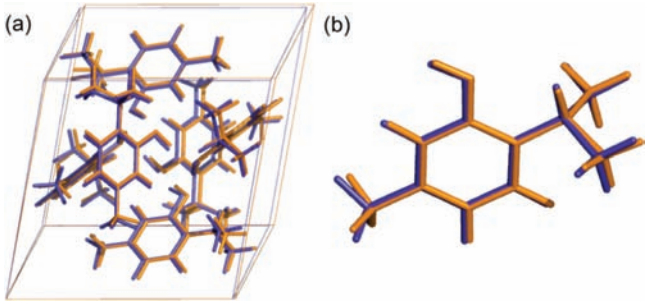

Figure 4. (a) Rhombohedral unit cell for the predicted structure determined to be in best agreement with the ^{1}H NMR observables (blue) and the known reference structure (orange). (b) Comparison of the two closest molecules in both structures.

In conclusion, agreement between calculated and experimental ^{1}H isotropic chemical shifts has been shown for the case of thymol to be sufficient criterion for the determination of the structure of a polycrystalline powder from *ab initio* predicted structures. We note that isotropic shifts are by far the easiest solid-state NMR parameter to measure. This method is very fast, assignment is not needed, and this approach should be applicable to sets of potential structures generated from a wide range of methods. The excellent results here have encouraged further work to test the general applicability of the method to molecular organic crystals, which is underway. Although they require a longer experiment, proton−proton spin-diffusion data are also shown to determine the correct structure and could prove useful when cross-validation of the structure determination is needed, for example if two potential candidates emerge from the chemical ^{1}H shift procedure.

Acknowledgment. Support from the ANR 06 Blanc PSD-NMR is gratefully acknowledged. G.M.D. thanks the Royal Society for funding via University Research Fellowship. We thank J.-N. Dumez for stimulating discussions and J. R. Yates for technical help with the CASTEP code.

Supporting Information Available: Three figures. Detailed CSP, CS and PSD calculation protocols. Predicted energies, experimental CS, CS rms errors, χ^2 and spacegroups for the 23 structures. Complete ref 1. Coordinates in CIF format and calculated CS. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Day, G. M.; et al. *Acta Crystallogr., Sect. B* **2009**, *65*, 107.
(2) (a) Payne, R. S.; Roberts, R. J.; Rowe, R. C.; Docherty, R. *Int. J. Pharm.* **1999**, *177*, 231. (b) Cabeza, A. J. C.; Day, G. M.; Motherwell, W. D. S.; Jones, W. *J. Am. Chem. Soc.* **2006**, *128*, 14466. (c) Panina, N.; Leusen, F. J. J.; Janssen, F.; Verwer, P.; Meekes, H.; Vlieg, E.; Deroover, G. *J. Appl. Crystallogr.* **2007**, *40*, 105.
(3) Beyer, T.; Day, G. M.; Price, S. L. *J. Am. Chem. Soc.* **2001**, *123*, 5086.
(4) (a) Bond, A. D.; Jones, W. *Acta Crystallogr., Sect. B* **2002**, *58*, 233. (b) Schmidt, M. U.; Ermrich, M.; Dinnebier, R. E. *Acta Crystallogr., Sect. B* **2005**, *61*, 37. (c) Day, G. M.; van de Streek, J.; Bonnet, A.; Burley, J. C.; Jones, W.; Motherwell, W. D. S. *Cryst. Growth Des.* **2006**, *6*, 2301. (d) Perrin, M. A.; Neumann, M. A.; Elmaleh, H.; Zaske, L. *Chem. Commun.* **2009**, 3181. (e) Fortes, A. D.; Suard, E.; Lemee-Cailleau, M. H.; Pickard, C. J.; Needs, R. J. *J. Am. Chem. Soc.* **2009**, *131*, 13508.
(5) Harris, R. K. *Solid State Sci.* **2004**, *6*, 1025.
(6) (a) Liu, F.; Orendt, A. M.; Alderman, D. W.; Grant, D. M. *J. Am. Chem. Soc.* **1997**, *119*, 8981. (b) Harper, J. K.; Grant, D. M. *J. Am. Chem. Soc.* **2000**, *122*, 3708. (c) Smith, E. D. L.; Hammond, R. B.; Jones, M. J.; Roberts, K. J.; Mitchell, J. B. O.; Price, S. L.; Harris, R. K.; Apperley, D. C.; Cherryman, J. C.; Docherty, R. *J. Phys. Chem. B* **2001**, *105*, 5818. (d) Rajeswaran, M.; Blanton, T. N.; Zumbulyadis, N.; Giesen, D. J.; Conesa-Moratilla, C.; Misture, S. T.; Stephens, P. W.; Huq, A. *J. Am. Chem. Soc.* **2002**, *124*, 14450. (e) Pickard, C. J.; Salager, E.; Pintacuda, G.; Elena, B.; Emsley, L. *J. Am. Chem. Soc.* **2007**, *129*, 8932. (f) Heider, E. M.; Harper, J. K.; Grant, D. M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 6083. (g) Seyfarth, L.; Sehnert, J.; El-Gamel, N. E. A.; Milius, W.; Kroke, E.; Breu, J.; Senker, J. *J. Mol. Struct.* **2008**, *889*, 217. (h) Barnes, A. B.; Andreas, L. B.; Huber, M.; Ramachandran, R.; van der Wel, P. C. A.; Veshtort, M.; Griffin, R. G.; Mehta, M. A. *J. Magn. Reson.* **2009**, *200*, 95.
(7) (a) Brouwer, D. H. *J. Am. Chem. Soc.* **2008**, *130*, 6306. (b) Cadars, S.; Lesage, A.; Pickard, C. J.; Sautet, P.; Emsley, L. *J. Phys. Chem. A* **2009**, *113*, 902. (c) Rossini, A. J.; Mills, R. W.; Briscoe, G. A.; Norton, E. L.; Geier, S. J.; Hung, I.; Zheng, S.; Autschbach, J.; Schurko, R. W. *J. Am. Chem. Soc.* **2009**, *131*, 3317. (d) Ashbrook, S. E.; Cutajar, M.; Griffin, J. M.; Lethbridge, Z. A. D.; Walton, R. I.; Wimperis, S. *J. Phys. Chem. C* **2009**, *113*, 10780. (e) Widdifield, C. M.; Bryce, D. L. *Phys. Chem. Chem. Phys.* **2009**, *11*, 7120. (f) Laurencin, D.; Gervais, C.; Wong, A.; Coelho, C.; Mauri, F.; Massiot, D.; Smith, M. E.; Bonhomme, C. *J. Am. Chem. Soc.* **2009**, *131*, 13430.
(8) Facelli, J. C.; Grant, D. M. *Nature* **1993**, *365*, 325.
(9) Harper, J. K.; Grant, D. M. *Cryst. Growth Des.* **2006**, *6*, 2315.
(10) (a) Brown, S. P.; Spiess, H. W. *Chem. Rev.* **2001**, *101*, 4125. (b) Goward, G. R.; Sebastiani, D.; Schnell, I.; Spiess, H. W.; Kim, H. D.; Ishida, H. *J. Am. Chem. Soc.* **2003**, *125*, 5792.
(11) Elena, B.; Emsley, L. *J. Am. Chem. Soc.* **2005**, *127*, 9140.
(12) Elena, B.; Pintacuda, G.; Mifsud, N.; Emsley, L. *J. Am. Chem. Soc.* **2006**, *128*, 9555.
(13) Salager, E.; Stein, R. S.; Pickard, C. J.; Elena, B.; Emsley, L. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2610.
(14) Day, G. M.; Motherwell, W. D. S.; Jones, W. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1693.
(15) Thozet, A.; Perrin, M. *Acta Crystallogr., Sect. B* **1980**, *36*, 1444.
(16) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. J.; Refson, K.; Payne, M. C. *Z. Kristallogr.* **2005**, *220*, 567.
(17) Pickard, C. J.; Mauri, F. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2001**, 63.
(18) Yates, J. R.; Pickard, C. J.; Mauri, F. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, 76.
(19) (a) Harris, R. K.; Hodgkinson, P.; Pickard, C. J.; Yates, J. R.; Zorin, V. *Magn. Reson. Chem.* **2007**, *45*, S174. (b) Johnston, J. C.; Iuliucci, R. J.; Facelli, J. C.; Fitzgerald, G.; Mueller, K. T. *J. Chem. Phys.* **2009**, *131*, 144503.

JA909449K