



HAL
open science

Reducing the computational cost of NMR crystallography of organic powders at natural isotopic abundance with the help of ^{13}C - ^{13}C dipolar couplings

Pierre Thureau, Simone Sturniolo, Miri Zilka, Fabio Ziarelli, Stéphane Viel,
Jonathan R Yates, Giulia Mollica

► To cite this version:

Pierre Thureau, Simone Sturniolo, Miri Zilka, Fabio Ziarelli, Stéphane Viel, et al.. Reducing the computational cost of NMR crystallography of organic powders at natural isotopic abundance with the help of ^{13}C - ^{13}C dipolar couplings. *Magnetic Resonance in Chemistry*, 2019, Special Issue "NMR Crystallography", 10.1002/mrc.4848 . hal-02067930

HAL Id: hal-02067930

<https://hal.science/hal-02067930>

Submitted on 14 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reducing the computational cost of NMR crystallography of organic powders at natural isotopic abundance with the help of ^{13}C - ^{13}C dipolar couplings

Pierre Thureau¹ | Simone Sturniolo² | Miri Zilka³ |
Fabio Ziarelli⁴ | Stéphane Viel^{1,5} | Jonathan R. Yates⁶
| Giulia Mollica^{1*}

¹Aix Marseille Univ, CNRS, ICR, Marseille, France

²Scientific Computing Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK

³Department of Physics, University of Warwick, Coventry CV4 7AL, UK

⁴Aix Marseille Univ, CNRS, Centrale Marseille, FSCM FR1739, Marseille, France

⁵Institut Universitaire de France, Paris, France

⁶Department of Materials, University of Oxford, Oxford OX1 3PH, UK

Correspondence

Giulia Mollica PhD, Aix Marseille Univ, CNRS, ICR, Marseille, France
Email: giulia.mollica@univ-amu.fr

Present address

[‡]Department, Institution, City, State or Province, Postal Code, Country

Funding information

European Research Council, Grant/Award

Number: 758498, EPSRC, Grant/Award

Number: EP/M022501/1

Abbreviations: CSP, crystal structure prediction; DQ, double quantum; AIRSS, *ab initio* random structure search; DFT, density functional theory; CSD, Cambridge structural database

Structure determination of functional organic compounds remains a formidable challenge when the sample exists as a powder. NMR crystallography approaches based on the comparison of experimental and DFT-computed ^1H chemical shifts have already demonstrated great potential for structure determination of organic powders, but limitations still persist. In this study we discuss the possibility of using ^{13}C - ^{13}C dipolar couplings quantified on powdered theophylline at natural isotopic abundance with the help of dynamic nuclear polarisation, to realize a DFT-free, rapid, screening of a pool of structures predicted by AIRSS. We show that while ^{13}C - ^{13}C dipolar couplings can identify structures possessing long-range structural motifs and unit cell-parameters close to those of the true structure, it must be complemented with other data to recover information about the presence and the chemical nature of the supramolecular interactions.

KEYWORDS

NMR, ^{13}C , dipolar coupling, crystal structure prediction, crystallography, natural abundance

1 | INTRODUCTION

The knowledge of the atomic-level structure of a material is fundamental to understand the origin of its end-use properties. Although single-crystal diffraction techniques have empowered the determination of three-dimensional molecular structure for a wide range of molecules and materials, diffraction techniques are much less efficient when the system under study does not form large crystals (>100 nm). Interestingly, recent progress in crystal structure prediction (CSP) methods demonstrated that crystal structures of powders can be computed *ab initio* with great accuracy.[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] CSP consists in exploring the crystal energy landscape of a molecule, *i.e.* finding "the set of crystal structures that are sufficiently low in energy to be thermodynamically plausible polymorphs".[2] Exploring the energy landscape of an organic molecule is computationally demanding, and its complexity depends on the molecule itself. For a given molecule, CSP usually generates a pool of candidate structures, which are ranked and then selected uniquely on the basis of their global lattice energy (instead of the more meaningful, but computationally more intensive, free energy). Therefore, structure ranking may prove difficult, especially in the case where the lattice energy difference between two candidate structures corresponds to the uncertainty of the CSP calculation (typically, a few kJ/mol).[13, 14, 2] This aspect is particularly important for organic molecular crystals, in which observed (or potentially observable) polymorphs typically correspond to different local minima on the energy landscape whose energy differs by only a few kJ/mol.[13] In this respect, it has been recently shown that the use of *ab initio* random structure search (AIRSS) method[15] can be beneficial because the random search algorithm ensures an unbiased sampling of the energy landscape, limiting the risk of missing energetically similar structures sitting far away on the energy landscape. For example in a recent study a set of 600 structures generated using the AIRSS method were used to identify two polymorphic forms of mABA.[16]

A solution to the structure-ranking problem is to combine CSP computation with experimental data, such as X-ray diffraction patterns (XRD) or data obtained from solid-state nuclear magnetic resonance (SSNMR) experiments[17, 18], to select the most relevant structure(s) from the pool of candidate structures generated by CSP.[19] For example, Xu *et al.* have combined CSP and X-ray diffraction to establish the crystal structure of glycine dihydrate.[20] Brus and coworkers showed that covariance between experimental and theoretical chemical shifts can be used to assign 2D heteronuclear correlation spectra and to refine the structure of organic solids by identifying the space group of the crystal in the absence of powder diffraction data.[21] Moreover, Baías *et al.* have used a combination of CSP and SSNMR to investigate the crystal structure of several organic molecules as powders.[22, 23] Notably, it has been shown that ^1H chemical shifts can be computed by density functional theory (DFT) for each candidate structure and that the comparison between computed and experimental chemical shifts could yield the correct structure.

In some cases, however, XRD and SSNMR cannot be exploited to assess the candidate structures generated by CSP. In fact, slight differences between a predicted structure and the correct structure may lead to important differences between experimental and calculated powder XRD data. [24, 16] Furthermore, SSNMR may be hampered by the low resolution often encountered for ^1H SSNMR spectra, the limited ^1H chemical shift range, and by the sparse number of distinguishable ^1H resonances. In addition, although chemical shift is certainly the most easily accessible NMR observable, its isotropic value is not directly related to the crystal structure. An additional limitation is that DFT computation of chemical shifts is a lengthy process, which further restricts the potential of SSNMR for structural

elucidation of organic powders.[22]

Consequently, there is a requirement for additional SSNMR observables to improve the validation or the refinement of CSP candidate structures. Interestingly, some of us have shown that double-quantum (DQ) experiments based on ^{13}C - ^{13}C dipolar couplings can yield quantitative information about the molecular conformation and the crystal packing of organic powders at natural abundance (NA).[25] The main advantage of this approach is that dipolar couplings are accessed through the corresponding dipolar curves, whose analytical expression is known from theory[26]. With the sole knowledge of the interatomic distances between the atoms in the crystal structure, dipolar curves can be straightforwardly calculated and plotted on a personal computer, with no need for first principle calculations. Therefore, an analytical function can be used to rapidly compare DQ ^{13}C - ^{13}C dipolar data obtained experimentally with those calculated for model crystal structures. Because of the low natural abundance of ^{13}C nuclear spins, the detection and measurement of ^{13}C - ^{13}C dipolar couplings in NA samples was only possible thanks to the optimal sensitivity enhancement provided by modern high-field dynamic nuclear polarization (DNP).[27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39] Importantly, DQ ^{13}C - ^{13}C dipolar data recorded on NA samples are directly related to short and long-range internuclear distances between the coupled spins because dipolar truncation can be neglected in the presence of naturally diluted ^{13}C - ^{13}C spin pairs.[40, 41, 42, 43]

In this contribution, we show that a pool of structures generated using AIRSS[15] can be screened at minimal computational cost with the help of DQ dipolar ^{13}C - ^{13}C build-up curves obtained by DNP SSNMR, demonstrating the sensitivity of these curves to a broad range of structural variations. We illustrate that this procedure can serve as a preliminary screening of the structures able to identify long range structural motifs that resemble those of the correct structure, allowing a coarse analysis of the structures *before* engaging in lengthy and computationally intensive structure refinement and chemical shift first principles calculations. The procedure is illustrated on theophylline, a molecular system that requires laborious crystallization work for XRD measurements,[44, 45, 46, 47] and for which the structural determination method based on the combination of CSP and ^1H chemical shifts failed.[22]

2 | MATERIALS AND METHODS

2.1 | Sample preparation

Powdered anhydrous theophylline was purchased from Sigma Aldrich and used as received. The powder was identified as being polymorph II of theophylline by both SSNMR and X-ray diffraction, corresponding to BAPLOT06 refcode in the crystallographic structural database (CSD). The sample for DNP analysis was prepared through incipient wetness impregnation[48] of the powder with a 66 mM solution of the polarising agent TEKPol[49] in 1,1,2,2-tetrabromoethane, as previously reported[25]. The wet powder was then transferred into a 3.2 mm sapphire rotor for DNP analysis. Absence of structural modifications induced by the preparation procedure was confirmed by comparing the SSNMR ^{13}C spectra and X-ray diffraction patterns collected before and after the impregnation.

2.2 | Acquisition of quantitative dipolar data by DNP NMR

All the DNP NMR experiments were performed using a Bruker 9.4 T wide bore magnet operated by a Avance-III console and equipped with a 3.2 mm low-temperature double-channel ^1H - ^{13}C , ^{29}Si CP-MAS probe. The sample temperature was 105 K and the spinning speed was 8 kHz. The NMR spectrometer was connected through a waveguide to a 263 GHz gyrotron delivering a micro-wave (MW) irradiation beam of 4 W at the probe. The field sweep coil of the NMR magnet was calibrated so that the MW irradiation of the gyrotron could match the maximum DNP enhancement obtained for

TOTAPOL (263.3 GHz).

The NMR pulse sequence used to obtain the double quantum (DQ) ^{13}C - ^{13}C dipolar build-up curves is shown in Fig. S4 (ESI). The POSTC7 scheme[50] was used to reintroduce the dipolar coupling interaction during both the excitation and the reconversion times τ_{DQ} of a 2D DQ correlation experiment. Each 2D experiment consisted of 34 t_1 increments, each with 64 acquisitions. The total evolution time for the indirect dimension was 0.6 ms, and the spectral width 56 kHz. States-TPPI procedure was applied to obtain pure-absorption spectra. A total of 9 2D DQ experiments were acquired by increasing the τ_{DQ} between 0.25 and 8 ms, each of which lasted 7 h (total experimental time: 63 h).

2.2.1 | Experimental NMR data

To obtain the DQ build-up curves, several 2D DQ experiments were acquired by increasing the τ_{DQ} , all the other parameters being unchanged. The volumes of the correlation peaks in each 2D spectrum (not shown here: see Ref.[25] for a full discussion of the data) were extracted and plotted as a function of τ_{DQ} , providing DQ build-up curves for each ^{13}C - ^{13}C pair of theophylline. The DQ signals were scaled with respect to both the number of ^{13}C sites contributing to the correlation peak analysed and the intensity of the corresponding correlation peaks in the CP-MAS spectrum acquired under equivalent experimental conditions (*i.e.* same nutation frequency and contact time). The confidence intervals for the DQ intensity measurements were determined by integrating several signal-free regions of the spectra of the same frequency bandwidths as those used to estimate the peak integrals. The standard deviation of these noise integrals were taken as the experimental confidence limits.

2.2.2 | Theoretical dipolar build-up curves

For a full derivation of the analytical function describing the evolution of DQ signal amplitudes under a γ -encoded pulse sequence, the reader is referred to Ref.[26]. Briefly, the theoretical DQ build-up curves for an isolated spin pair interacting through dipolar coupling is the following:

$$s_{ij}(\tau_{DQ}) = \frac{1}{2} - \frac{1}{x\sqrt{8}} \left[F_c(x\sqrt{2})\cos 2\theta + F_s(x\sqrt{2})\sin 2\theta \right] \quad (1)$$

where:

$$x = \frac{2\theta}{\pi} \quad (2)$$

$$\theta = \frac{3}{2} k_{C7} b_{ij} \tau_{DQ} \quad (3)$$

In Eq. 1, F_c and F_s represent the cosine and sine Fresnel integrals, respectively, while in Eq. 3, k_{C7} indicates the POSTC7 scaling factor, equal to 0.155, and b_{ij} is the dipolar coupling constant between the spins i and j , which depends on their internuclear distance.

The analytical function used to mimic the DQ evolution of a carbon atom pair ij in a crystal has the following form:

$$S_{ij}(\tau_{DQ}) = A \exp(-2k \cdot \tau_{DQ}) \sum_{j=1}^{N_j} s_{ij}(\tau_{DQ}) \quad (4)$$

Importantly, Eq. 4 only requires the knowledge of the DQ recoupling scaling factor and the list of internuclear distances between the sites i and j in the crystal structure to allow the theoretical dipolar build-up curves to be calculated for this atom pair. The factor A defines the intensity of DQ signal, while k accounts for relaxation effects.

Relying on a purely analytical function, the calculation of the DQ build-up curves is immediate and allows a rapid screening of a large number of structures. For each crystal structure, the theoretical curves were obtained from Eq. 4 by taking into account all the internuclear distances between two chosen carbon atoms in the crystal up to a maximum distance of 6.5 Å. In the case of theophylline, including distances longer than 6.5 Å did not bring any significant modification to the curves. Importantly, these curves receive contributions from both intra- and intermolecular dipolar interactions: therefore, although they cannot be interpreted in terms of a single, specific, internuclear distance, they encode the crystal structure.

2.2.3 | Dipolar χ^2 calculation

For each computed structure, the ^{13}C - ^{13}C dipolar χ^2 for the atom pair ij was calculated as:

$$\chi_{ij}^2 = \sum_{\tau_{DQ}} \left[S_{ij}^{\text{exp}}(\tau_{DQ}) - S_{ij}(\tau_{DQ}) \right]^2 \quad (5)$$

where the experimental signal amplitude $S_{ij}^{\text{exp}}(\tau_{DQ})$ was obtained as described in Section 2.2. The global χ^2 discussed in the text was obtained by adding up the χ_{ij}^2 for the 8 distinct ij atom pairs listed in Table S2 (ESI).

2.3 | Crystal structure prediction

A set of 33 candidate crystal structures was generated using the Ab Initio Random Structure Search (AIRSS) approach [51, 15]. Structure prediction using AIRSS involves two stages: trial structures are first generated through the random search component, and then they are geometry optimised through a fully periodic DFT calculation (see next section). For theophylline, the first stage of calculation used neutral theophylline molecules with fixed conformation taken from the X-ray determined structure (CSD entry code: BAPLOT06). All the structures were generated by constraining the space group to Pn21, corresponding to the space group observed for the CSD structure of theophylline. Therefore, all the structures had 4 molecules in the unit cell, whose dimensions were allowed to vary.

No constraints were applied to the unit-cell, ensuring that the resulting structures have different long-range arrangement in terms of mutual molecular distance and orientation. It is worth noting that, except for the reorientation of the methyl groups about their ternary axis, the (planar) molecule of theophylline does not have any degrees of freedom, therefore only minimal conformational differences are observed among the predicted structures. It follows that all the candidate structures predicted by AIRSS are pure packing polymorphs of theophylline.

2.4 | First-principles calculations

Geometry optimization of the structures predicted by AIRSS was carried out using the DFT code CASTEP,[52, 53, 54, 55] which uses a plane wave basis set together with pseudo potentials to represent the core-valence interaction. All calculations used the PBE functional[56] and the dispersion correction scheme of Tkatchenko-Scheffler[57]. Calculations carried out on trial structures generated in the first stage of AIRSS search used a plane wave cut-off energy of 500 eV and a Brillouin zone sampling of $0.1 \times 2\pi \text{ \AA}^{-1}$. The convergence criteria were 0.05 eV \AA^{-1} for forces, 0.1 GPa for stresses, 0.00002 eV per atom for energy and 0.001 \AA for atomic displacements. Although the resulting structures are energetically stable, and can hence be considered physically meaningful, their energy lie within 60 kJ/mol from the lowest energy structure. These values are higher than the threshold of 10 kJ/mol typically reported in NMR crystallography studies. However, the scope of the present work was not to solve the structure of theophylline, but to propose a new approach to rapidly screen a set of physically meaningful candidate crystal structures. Therefore, no particular effort was put into obtaining refined structures that could resemble the true structure of theophylline. The lattice energies of each structure relative to the energy of the known crystal structure (referred to as CSD in the following) are reported in Table S1 and Fig. S1 (ESI).

3 | RESULTS AND DISCUSSION

The trial set of 33 candidate structures of anhydrous theophylline, generated by the AIRSS method, are indicated in the following with the corresponding number (1 to 33). In addition, the CSD structure is used here as a structural reference. Structure numbering does not reflect increasing lattice energy. Because of the minimal conformational rearrangement allowed for theophylline molecules, variations in the ^{13}C - ^{13}C dipolar curves between different predicted structures should be uniquely ascribed to differences in the mutual position and/or orientation of the - rigid - theophylline molecules, and not to intramolecular differences. Considering that dipolar couplings scale as $1/r^3$, intermolecular contributions are typically 1 order of magnitude smaller than intramolecular contributions ($\sim 100 \text{ Hz}$ vs $\sim 2000 \text{ Hz}$). It follows that theophylline is an appropriate, yet challenging, system to probe the ability of such small variations in ^{13}C - ^{13}C dipolar couplings to provide a reliable tool to assess the quality of crystal structure candidates.

Fig. 1 shows experimental ^{13}C - ^{13}C DQ dipolar build-up curves measured by MAS DNP for different carbon atom pairs of NA powdered anhydrous theophylline [25]. These data are compared with the theoretical build-up curves calculated for some of the candidate structures predicted by AIRSS. Specifically, in Fig. 1 the dipolar build-up curves calculated for the CSD structure are only compared with the curves for structures n.12 and 21 (panels b, d and f), and for structures n.32 and 33 (panels c, d and g) for the sake of clarity. The C-C pairs shown in Fig.1 are representative of the variability of information contained in ^{13}C - ^{13}C dipolar curves. For instance, while the C1-C6,7 pair (Fig.1b and c) receives contributions from both intra and intermolecular couplings, only intermolecular interactions contribute to the C5-C5 pair (Fig.1d and e). For all the spin pairs considered, the curves calculated for the CSD structure are in very good agreement with the experimental data. Interestingly, the dipolar curves plotted for the C1-C6,7 pair in Fig.1b and c are in good agreement with the experimental data for the four predicted structures n.12, 21, 32 and 33. Nonetheless, Fig.1 also confirms that different crystal structures can have dipolar curves that almost superpose for certain atom pairs but differ significantly for others. This is illustrated in Fig.1d and e, showing that the intermolecular C5-C5 distances are in agreement with the experimental data for structures n.12, 32 and 33, but not for n.21. The case of C6-C7 pair is particularly worth discussing. In fact, although the dipolar curves for C6-C7 are built from both intra and intermolecular contacts, in most of the candidate structures of theophylline considered in this analysis (including the case of the CSD structure) we observed that the shortest intermolecular contribution to the C6-C7 coupling was

stronger than the intramolecular one. Dipolar data corresponding to atom pairs such as the C6-C7 of theophylline are hence expected to show high sensitivity to the crystal packing and are very precious for structure determination purposes. This is illustrated in Fig.1f and g, showing that dipolar build-up curves calculated for the C6-C7 pair varied sensibly from one structure to the other, and only agrees with experimental data for structures n.32 and 33. Importantly,

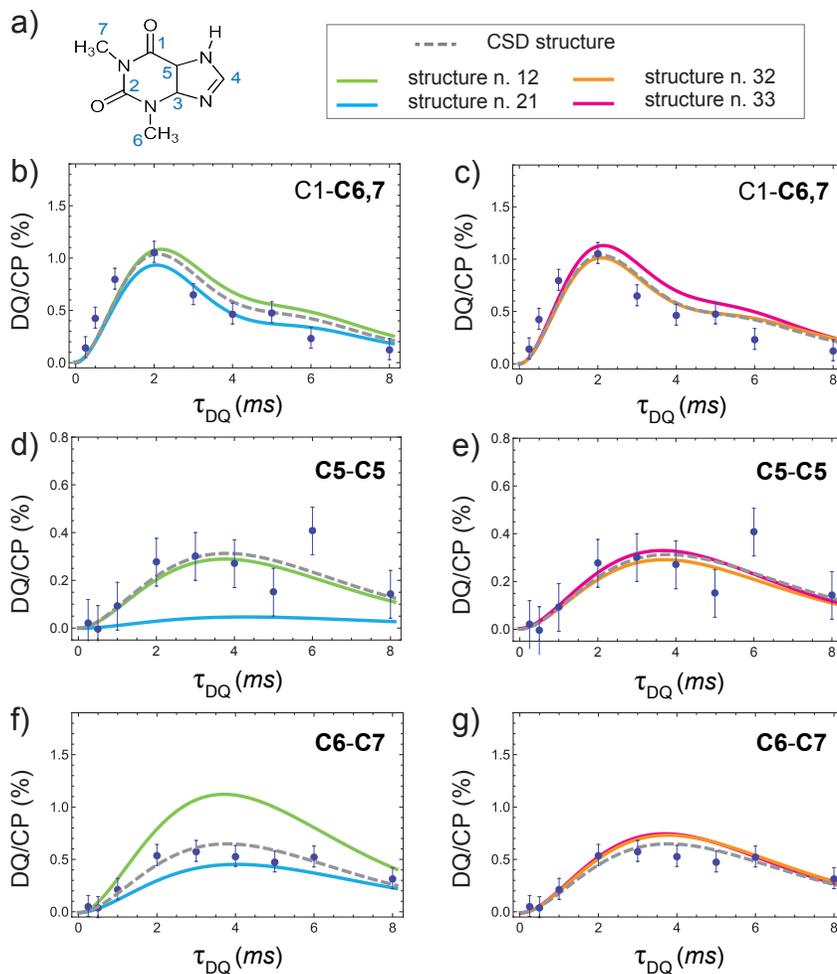


FIGURE 1 DQ build-up curves for selected pairs of ^{13}C dipolar correlation peaks. Selected pairs of peaks are indicated by numbers on the top right of each graph, with the number in bold referring to the detected resonance. Labeling refers to the scheme in a). The experimental data points (blue plain circles) were obtained by integrating the corresponding correlation peaks obtained in a series of DNP-enhanced 2D ^{13}C - ^{13}C DQ correlation spectra (see Ref. [25] and ESI) recorded for different τ_{DQ} values on anhydrous theophylline. The solid lines represent the analytical function of the DQ build-up, which depends on the ^{13}C - ^{13}C interatomic distances in structures: n.12 and 21 (b, d and f) and n.32 and 33 (c, e and g). The curve corresponding to the CSD structure is also shown as a dashed grey line. Evaluation of the deviation of each of these curves from the experimental data points (χ^2) allows easy screening of the whole set of trial structures at no computational cost.

for this atom pair, the best agreement between experimental data and calculated curves was neatly observed for the CSD structure. The method is hence sensitive to recurrent supramolecular motifs in molecular crystals, as it happens for $\pi - \pi$ stacking in the case of theophylline.

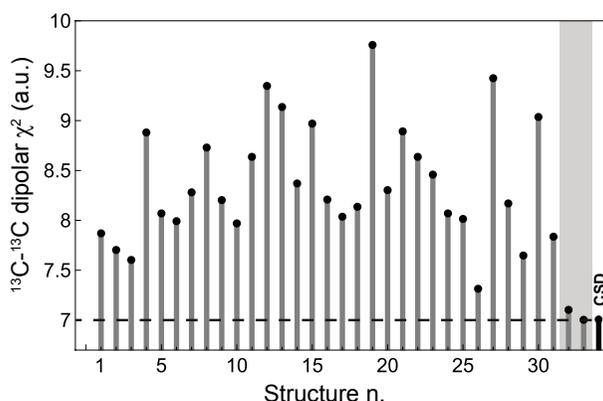


FIGURE 2 ^{13}C - ^{13}C global dipolar χ^2 for the crystal structures of theophylline generated using AIRSS. The χ^2 for the CSD structure is also shown for comparison. The dotted line is a guide to the eyes to identify the χ^2 minimum. The predicted structures showing the best agreement are highlighted in gray (n.32 and 33).

The deviation between the experimental and theoretical ^{13}C - ^{13}C DQ dipolar build-up curves can be used to quantify the goodness of a candidate structure. As described in section 2.2.3, such deviation can be expressed as the global dipolar χ^2 , reported in Fig. 2 for all the predicted structures and the CSD structure. This value is obtained by adding the dipolar χ^2 obtained for 8 different ^{13}C - ^{13}C pairs (see Table S2, ESI). The results show that the best agreement is obtained for AIRSS structures n. 32 and 33. It must be noted that, differently from the RMSD typically used for evaluating the goodness of a structure based on the difference between calculated and experimental chemical shifts, χ^2 takes into account the experimental error in the determination of the dipolar coupling. Although a confidence interval for χ^2 values cannot be established using a set of dipolar data obtained on a single compound, structures n. 32 and 33 can be retained as good candidate structures in the present analysis because they show nearly identical values of χ^2 , which essentially coincide with the χ^2 for the CSD structure.

Importantly, the best agreement between the experimental and theoretical ^{13}C - ^{13}C DQ dipolar build-up curves is obtained for two candidate structures of theophylline, n.32 and 33, that show structural features close to the CSD structure. In fact, as shown in Fig. 3, theophylline molecules in these structures have very similar supramolecular arrangement (i.e. similar mutual arrangement, similar location in the unit cell) but differ in their mutual *orientation*, leading to different intermolecular interactions. Notably, while in the CSD structure the amino proton is hydrogen bonded to the aldimine nitrogen ($\text{NH} \cdots \text{N}$), the amino proton in structure n.33 is hydrogen bonded to the carbonyl oxygen neighboring the imidazole ring ($\text{NH} \cdots \text{O}$).

It is also interesting to consider the similarities of the selected structures in terms of their unit-cell parameters. As shown in Fig. 4, three clusters of structures can be identified considering their unit-cell values. The distribution of the structures in the three clusters does not follow an energy criterium, since both high and low energy structures populate each of the clusters (Fig. S3, ESI). However, Fig. 4 shows that the unit cell parameters of structures n. 32 and 33 are nearly identical to those of the CSD structure. Globally, this analysis demonstrates that ^{13}C dipolar data are sensitive to both unit-cell dimensions and supramolecular arrangement. However, when carbon diluted, planar,

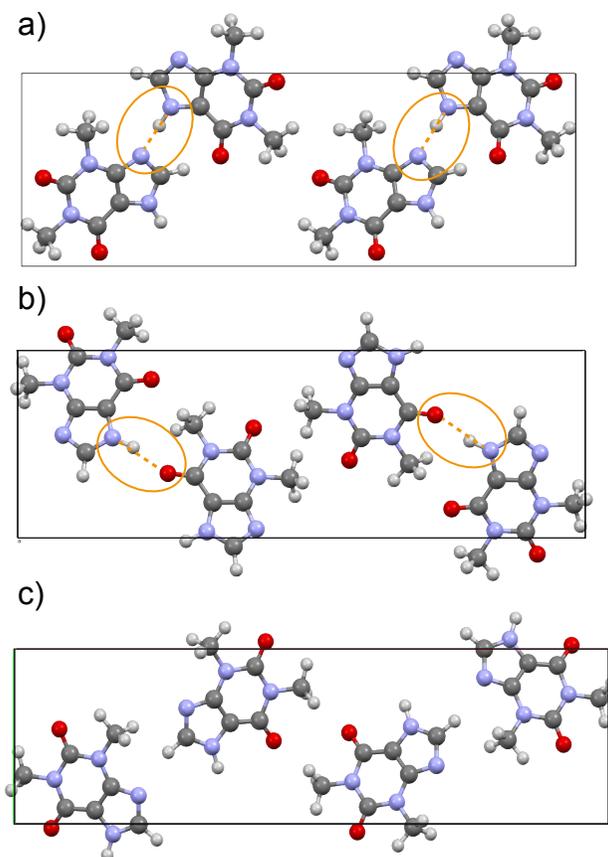


FIGURE 3 View down one of the crystallographic axes of three candidate crystal structures of theophylline. Structure a) corresponds to CSD structure and coincides with the structure deposited in the CSD. This structure is stabilized by the presence of N-H...N hydrogen bonds. In b), corresponding to predicted structure n.33, this pattern is substituted by N-H...O hydrogen bonds. These two crystal structures are virtually indistinguishable on the basis of their global dipolar χ^2 . Despite being very similar to a), structure c), corresponding to predicted structure n.32, does not allow the formation of hydrogen bonds: this structure is characterized by slightly larger dipolar χ^2 than a) and b). Crystal structures n.12 and 21 are also shown in Fig. S2 (ESI) for comparison.

molecules such as theophylline are investigated, ^{13}C dipolar data appear not sensitive enough to efficiently discriminate between structures only differing by in-plane rotation of one of the molecules.[58]

Nonetheless, using dipolar data, the pool of AIRSS structures of theophylline examined in this study could be assessed within a few seconds on a standard computer, without the need of first principle calculations, *i.e.* at essentially no computational cost. Contrastingly, computation of chemical shifts from first principles would have required about 1.5 CPU hours per structure on 12 cores, plus the time necessary for full geometry optimization. Differently from previously proposed methods, the approach proposed here does not rank structures based on their lattice energy, but relies on purely structural parameters (*i.e.* the internuclear distances in the crystal) to identify likely candidate structures. It is important to point out that ^{13}C dipolar data such as those analysed in the present study does not contain any chemical information, such as the presence/type of chemical bonds or the occurrence of interatomic

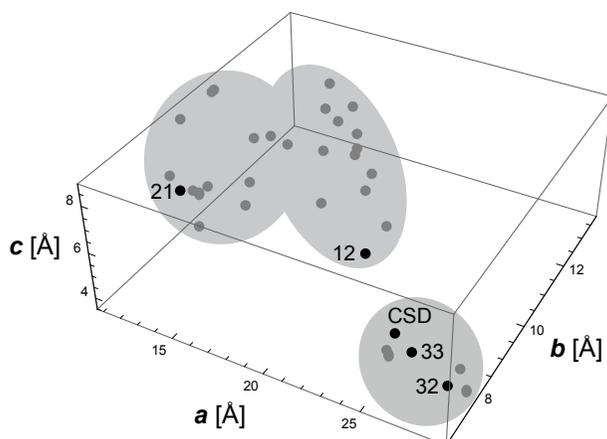


FIGURE 4 Unit-cell parameters for structures resulting from the AIRSS calculation on theophylline. Three clusters of structures can be identified based on the values of the unit-cell axes, indicated by the grey areas. Unit-cell parameters of structures n. 32 and 33, selected on the basis of dipolar data, are very similar to those of the CSD structure.

interactions. Notwithstanding these limitations, dipolar data helped identify a sub-set of candidate structures close to the global minimum without any preliminary structural assumption, reducing the initial pool of candidate structures of theophylline by more than 90%. The proposed method might be of interest in the perspective of investigating larger pools of candidate structures, for which NMR crystallography approaches based on chemical shifts are time consuming and computationally intensive. We suggest that, in the presence of large pools of properly geometry-optimized structures, calculation of chemical shifts could be restrained to the sub-set of structures identified by dipolar data, as a structure refinement. Therefore, the combination of dipolar data and chemical shifts could constitute a new protocol for faster and reliable structure determination of organic compounds at natural isotopic abundance that could considerably simplify the analysis of large sets of trial crystal structures. This work is currently in progress in our group and will make the object of a future study.

4 | CONCLUSIONS

The results presented here suggest that ^{13}C - ^{13}C DQ dipolar build-up curves acquired at natural isotopic abundance using DNP SSNMR might be used to sort out a pool of predicted crystal structures within a few seconds. This procedure is complementary to approaches based on NMR chemical shifts and PXRD experiments and a combination of all these techniques should allow in the future to solve the structure of unknown organic compounds at natural isotopic abundance. It should be noted that ^{13}C - ^{13}C DQ dipolar experiments can also be used to establish ^{13}C - ^{13}C connectivities,^[59] which ensures that the chemical structure of the predicted structures is reliable. In this form, we believe that DNP SSNMR experiments performed on natural abundance samples will be helpful to investigate the structure of powders that cannot be characterized using standard techniques.

ACKNOWLEDGEMENTS

GM acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 758498). First principles calculations were performed using the HPC resources of Aix-Marseille Université financed by the project Equip@Meso (ANR-10-EQPX-29-01) of the program "Investissements d'Avenir" supervised by the Agence Nationale de la Recherche. JRY and SS are supported by the Collaborative Computational Project for NMR Crystallography funded by the EPSRC (UK) Grant No. EP/M022501/1.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ENDNOTES

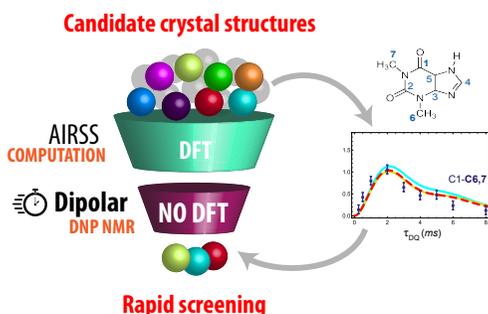
REFERENCES

- [1] Price, S. L. *Acc. Chem. Res.* **2009**, *42*, 117-126.
- [2] Price, S. L. *Chem. Soc. Rev.* **2014**, *43*, 2098-2111.
- [3] Reilly, A. M. *et al. Acta Crystallographica Section B* **2016**, *72*, 439-459.
- [4] Lyakhov, A. O.; Oganov, A. R.; Stokes, H. T.; Zhu, Q. *Comput. Phys. Commun.* **2013**, *184*, 1172-1182.
- [5] Day, G. M.; S. Motherwell, W. D.; Jones, W. *Physical Chemistry Chemical Physics* **2007**, *9*, 1693-1704.
- [6] Neumann, M.; Leusen, F.; Kendrick, J. *Angewandte Chemie International Edition* **2008**, *47*, 2427-2430.
- [7] Neumann, M. A. *The Journal of Physical Chemistry B* **2008**, *112*, 9810-9829.
- [8] Kendrick, J.; Leusen, F. J.; Neumann, M. A.; van de Streek, J. *Chemistry - A European Journal* **2011**, *17*, 10736-10744.
- [9] van de Streek, J.; Neumann, M. A. *CrystEngComm* **2011**, *13*, 7135-7142.
- [10] Day, G. M. *Crystallography Reviews* **2011**, *17*, 3-52.
- [11] Price, S. L. *Acta Crystallographica Section B* **2013**, *69*, 313-328.
- [12] Price, S. L.; Braun, D. E.; Reutzel-Edens, S. M. *Chem. Commun.* **2016**, *52*, 7065-7077.
- [13] Beran, G. J. O. *Chem. Rev.* **2016**, *116*, 5567-5613 PMID: 27008426.
- [14] Nyman, J.; Day, G. M. *CrystEngComm* **2015**, *17*, 5154-5165.
- [15] Pickard, C. J.; Needs, R. J. *Phys. Rev. Lett.* **2006**, *97*, 045504.
- [16] Zilka, M.; Dudenko, D. V.; Hughes, C. E.; Williams, P. A.; Sturniolo, S.; Franks, W. T.; Pickard, C. J.; Yates, J. R.; Harris, K. D. M.; Brown, S. P. *Phys. Chem. Chem. Phys.* **2017**, *19*, 25949-25960.
- [17] Martineau, C.; Senker, J.; Taulelle, F. Chapter One - NMR Crystallography. In , Vol. 82; Webb, G. A., Ed.; Academic Press: 2014.

- [18] Bonhomme, C.; Gervais, C.; Babonneau, F.; Coelho, C.; Pourpoint, F.; Azais, T.; Ashbrook, S. E.; Griffin, J. M.; Yates, J. R.; Mauri, F.; Pickard, C. J. *Chem. Rev.* **2012**, *112*, 5733-5779.
- [19] Dudek, M. K.; Bujacz, G.; Potrzebowski, M. *J. CrystEngComm* **2017**, *19*, 2903-2913.
- [20] Xu, W.; Zhu, Q.; Hu, C. T. *Angew. Chem. Int. Ed.* **2017**, *56*, 2030-2034.
- [21] Brus, J.; Czernek, J.; Kobera, L.; Urbanova, M.; Abbrent, S.; Husak, M. *Crystal Growth & Design* **2016**, *16*, 7102-7111.
- [22] Baias, M.; Widdifield, C. M.; Dumez, J.-N.; Thompson, H. P. G.; Cooper, T. G.; Salager, E.; Bassil, S.; Stein, R. S.; Lesage, A.; Day, G. M.; Emsley, L. *Phys. Chem. Chem. Phys.* **2013**, *15*, 8069-8080.
- [23] Baias, M.; Dumez, J.-N.; Svensson, P. H.; Schantz, S.; Day, G. M.; Emsley, L. *J. Am. Chem. Soc.* **2013**, *135*, 17501-17507.
- [24] Leclaire, J.; Poisson, G.; Ziarelli, F.; Pepe, G.; Fotiadu, F.; Paruzzo, F. M.; Rossini, A. J.; Dumez, J.-N.; Elena-Herrmann, B.; Emsley, L. *Chem. Sci.* **2016**, *7*, 4379-4390.
- [25] Mollica, G.; Dekhil, M.; Ziarelli, F.; Thureau, P.; Viel, S. *Angew. Chem. Int. Ed.* **2015**, *54*, 6028-6031.
- [26] Pileio, G.; Concistrè, M.; McLean, N.; Gansmüller, A.; Brown, R. C.; Levitt, M. H. *J. Magn. Reson.* **2007**, *186*, 65-74.
- [27] Maly, T.; Debelouchina, G. T.; Bajaj, V. S.; Hu, K.-N.; Joo, C.-G.; Mak-Jurkauskas, M. L.; Sirigiri, J. R.; van der Wel, P. C. A.; Herzfeld, J.; Temkin, R. J.; Griffin, R. G. *J. Chem. Phys.* **2008**, *128*, 052211.
- [28] Ni, Q. Z.; Daviso, E.; Can, T. V.; Markhasin, E.; Jawla, S. K.; Swager, T. M.; Temkin, R. J.; Herzfeld, J.; Griffin, R. G. *Acc. Chem. Res.* **2013**, *46*, 1933-1941.
- [29] Rossini, A. J.; Zagdoun, A.; Lelli, M.; Lesage, A.; Copéret, C.; Emsley, L. *Acc. Chem. Res.* **2013**, *46*, 1942-1951 PMID: 23517009.
- [30] Lilly Thankamony, A. S.; Lion, C.; Pourpoint, F.; Singh, B.; Perez Linde, A. J.; Carnevale, D.; Bodenhausen, G.; Vezin, H.; Lafon, O.; Polshettiwar, V. *Angew. Chem. Int. Ed.* **2015**, *54*, 2190-2193.
- [31] Smith, A. N.; Long, J. R. *Anal. Chem.* **2016**, *88*, 122-132.
- [32] Geiger, Y.; Gottlieb, H. E.; Akbey, Ü.; Oschkinat, H.; Goobes, G. *J. Am. Chem. Soc.* **2016**, *138*, 5561-5567 PMID: 26451953.
- [33] Hoffmann, M. M.; Bothe, S.; Gutmann, T.; Buntkowsky, G. *J. Phys. Chem. C* **2017**, *121*, 22948-22957.
- [34] Thankamony, A. S. L.; Wittmann, J. J.; Kaushik, M.; Corzilius, B. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *102-103*, 120-195.
- [35] Kobayashi, T.; Perras, F. A.; Murphy, A.; Yao, Y.; Catalano, J.; Centeno, S. A.; Dybowski, C.; Zumbulyadis, N.; Pruski, M. *Dalton Trans.* **2017**, *46*, 3535-3540.
- [36] Mentink-Vigier, F.; Vega, S.; De Paepe, G. *Phys. Chem. Chem. Phys.* **2017**, *19*, 3506-3522.
- [37] Mollica, G.; Le, D.; Ziarelli, F.; Casano, G.; Ouari, O.; Phan, T. N. T.; Aussenac, F.; Thureau, P.; Gigmes, D.; Tordo, P.; Viel, S. *ACS Macro Letters* **2014**, *3*, 922-925.
- [38] Marker, K.; Paul, S.; Fernandez-de Alba, C.; Lee, D.; Mouesca, J.-M.; Hediger, S.; De Paepe, G. *Chem. Sci.* **2017**, *8*, 974-987.
- [39] Rossini, A. J. *The Journal of Physical Chemistry Letters* **2018**, *9*, 5150-5159.
- [40] Saalwächter, K. *ChemPhysChem* **2013**, *14*, 3000-3014.
- [41] De Paëpe, G. *Annu. Rev. Phys. Chem.* **2012**, *63*, 661-684.

- [42] Ladizhansky, V. *Solid State Nuclear Magnetic Resonance* **2009**, *36*, 119–128.
- [43] Mollica, G.; Madhu, P. K.; Ziarelli, F.; Thévand, A.; Thureau, P.; Viel, S. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4359–4364.
- [44] Ebisuzaki, Y.; Boyle, P.; Smith, J. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1997**, *53*, 777.
- [45] Fucke, K.; McIntyre, G. J.; Wilkinson, C.; Henry, M.; Howard, J. A. K.; Steed, J. W. *Cryst. Growth Des.* **2012**, *12*, 1395–1401.
- [46] 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*, 5818–5826.
- [47] Pinon, A. C.; Rossini, A. J.; Widdifield, C. M.; Gajan, D.; Emsley, L. *Mol. Pharmaceut.* **2015**, *12*, 4146–4153 PMID: 26393368.
- [48] Lesage, A.; Lelli, M.; Gajan, D.; Caporini, M. A.; Vitzthum, V.; Miéville, P.; Alauzun, J.; Roussey, A.; Thieuleux, C.; Mehdi, A.; Bodenhausen, G.; Coperet, C.; Emsley, L. *J. Am. Chem. Soc.* **2010**, *132*, 15459–15461.
- [49] Zagdoun, A.; Casano, G.; Ouari, O.; Schwarzwälder, M.; Rossini, A. J.; Aussenac, F.; Yulikov, M.; Jeschke, G.; Copéret, C.; Lesage, A.; Tordo, P.; Emsley, L. *J. Am. Chem. Soc.* **2013**, *135*, 12790–12797.
- [50] Hohwy, M.; Jakobsen, H. J.; Edén, M.; Levitt, M. H.; Nielsen, N. C. *J. Chem. Phys.* **1998**, *108*, 2686–2694.
- [51] Pickard, C. J.; Needs, R. J. *J. Phys.: Condens. Matter* **2011**, *23*, 053201.
- [52] Segall, M. D.; Lindan, P. J. D.; Probert, M. J.; Pickard, C. J.; Hasnip, P. J.; Clark, S. J.; Payne, M. C. *J. Phys.: Condens. Matter* **2002**, *14*, 2717.
- [53] Pickard, C. J.; Mauri, F. *Phys. Rev. B* **2001**, *63*, 245101.
- [54] Yates, J. R.; Pickard, C. J.; Mauri, F. *Phys. Rev. B* **2007**, *76*, 024401.
- [55] Harris, R. K.; Hodgkinson, P.; Pickard, C. J.; Yates, J. R.; Zorin, V. *Magn. Reson. Chem.* **2007**, *45*, S174–S186.
- [56] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [57] Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.
- [58] Mollica, G.; Dekhil, M.; Ziarelli, F.; Thureau, P.; Viel, S. *Solid State Nucl. Magn. Reson.* **2015**, *65*, 114–121 {NMR} Crystallography.
- [59] Dekhil, M.; Mollica, G.; Bonniot, T. T.; Ziarelli, F.; Thureau, P.; Viel, S. *Chem. Commun.* **2016**, *52*, 8565–8568.

GRAPHICAL ABSTRACT



- A new DFT-free NMR crystallography approach is presented, that allows a pool of predicted structures to be screened using ¹³C-¹³C DQ dipolar curves measured at natural abundance with the help of DNP
- Tested on theophylline, the method was able to identify structures possessing long-range structural motifs and unit-cell parameters similar to those of the CSD structure, without any structural assumption
- This approach is expected to considerably simplify the analysis of large structure pools with respect to current computationally-

intensive approaches