

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/159960>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

© 2021 Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# A toolbox for improving the workflow of NMR crystallography

Patrick M.J. Szell<sup>1</sup>, Sten O. Nilsson Lill<sup>2</sup>, Helen Blade<sup>3</sup>, Steven P. Brown<sup>1\*</sup>, Leslie P.  
Hughes<sup>3\*</sup>

\*Author to whom correspondence may be addressed

<sup>1</sup> Department of Physics, University of Warwick, Coventry, UK, CV4 7AL. E-mail: [S.P.Brown@warwick.ac.uk](mailto:S.P.Brown@warwick.ac.uk)

<sup>2</sup> Early Product Development and Manufacturing, Pharmaceutical Sciences, R&D, AstraZeneca, Gothenburg, Sweden.

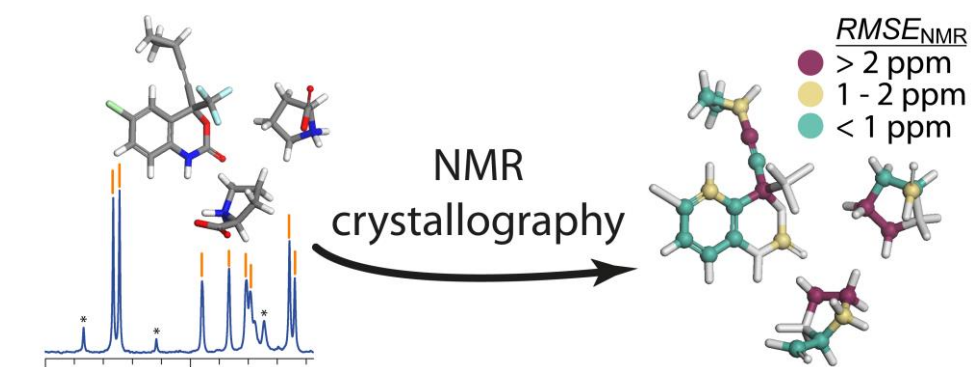
<sup>3</sup> Oral Product Development, Pharmaceutical Technology & Development, Operations, AstraZeneca, Macclesfield, UK. E-mail: [Les.Hughes2@astrazeneca.com](mailto:Les.Hughes2@astrazeneca.com)

**Abstract.**

NMR crystallography is a powerful tool with applications in structural characterization and crystal structure verification, to name two. However, applying this tool presents several challenges, especially for industrial users, in terms of consistency, workflow, time consumption, and the requirement for a high level of understanding of experimental solid-state NMR and GIPAW-DFT calculations. Here, we have developed a series of fully parameterized scripts for use in Materials Studio and TopSpin, based on the .magres file format, with a focus on organic molecules (e.g. pharmaceuticals), improving efficiency, robustness, and workflow. We separate these tools into three major categories: performing the DFT calculations, extracting & visualizing the results, and crystallographic modelling. These scripts will rapidly submit fully parameterized CASTEP jobs, extract data from the calculations, assist in visualizing the results, and expedite the process of structural modelling. Accompanied with these tools is a description on their functionality, documentation on how to get started and use the scripts, and links to video tutorials for guiding new users. Through the use of these tools, we hope to facilitate NMR crystallography and to harmonize the process across users.

**Keywords.** NMR crystallography, density functional theory, crystallographic disorder, solid-state NMR, automation

## Graphical abstract



## Highlight

- Automating NMR crystallography.
- Harmonizing the process of NMR crystallography across users.
- Accelerating data extraction, visualization, and structural modelling.

## 1 - Introduction

Solid-state NMR of organic molecules, e.g. pharmaceuticals, is useful for structural characterization, structural verification, and quality control, to name a few of its multiple applications.[1-25] Computational methods have been developed to calculate NMR observables, such as the chemical shifts, quadrupolar coupling, or  $J$ -coupling, with the implementation of periodicity facilitating the application to periodic solid-state structures. Gauge-including projector augmented wave (GIPAW)[26-28] density functional theory (DFT), as implemented in, e.g., CASTEP[29] and Quantum Espresso,[30-32] has been shown to reproduce chemical shifts typically within a 1% accuracy of the typical chemical-shift range, representing a discrepancy of 2 ppm for  $^{13}\text{C}$  relative to experiment.[28, 33-38] Solid-state NMR is complementary to other solid-state characterization methods, notably diffraction. The combination of experimental solid-state NMR and DFT calculations has been termed *NMR crystallography*,[7, 21, 24, 36, 39-57] finding application, for example, in crystallographic refinement.

Solid-state NMR chemical shifts can provide key crystallographic details, notably the number of molecules in the asymmetric unit ( $Z'$ ),[46, 58-61] thus, for example, facilitating the choice of candidate structures from crystal structure predictions (CSP)[49, 52, 62-64] and allowing crystal structure refinement using NMR observables.[49, 65-69] In addition,  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  NMR chemical shifts are known to be influenced by non-covalent interactions, e.g. hydrogen bonding and ring current effects,[12, 70-74] proton positions & protonation sites,[75-77] and can be useful for characterizing crystallographic disorder.[2, 60, 78-82] As a result, the analysis of  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  chemical shifts has been used to address key chemical challenges in organic molecules, such as determining whether a compound is a salt or a cocrystal.[6, 75, 83] While noting this wide breadth of applicability, in this paper we limit our discussion to structural refinement and validation using NMR chemical shifts.

Chemists interested in solids have much to benefit from the current developments in NMR crystallography. However, NMR crystallography can be difficult to adopt, especially in industry, in part due to the requirements of highly specialized knowledge and the time taken to implement the analysis. In addition, multiple parameters can be selected when performing NMR crystallography calculations, introducing at times a variability in the resulting output. These variations can lead to inconsistent results and can take time to correct, one example being that there is often different atom labelling between the input structural model and the CASTEP output. In an effort to ease the adoption of NMR crystallography, purpose-made tools are required, and ideally these tools would be consolidated within industry-standard software, e.g., *Materials Studio* (Biovia)[84] and *TopSpin* (Bruker),[85] in order to facilitate a consistent data workflow, but while making available the script codes to allow user adaption for more general applicability. Here, we report a series of computational tools to aid in the process of NMR crystallography, and we separate these tools into three major categories: performing the DFT calculations, extracting & visualizing the results, and crystallographic modelling. In order to standardise their implementation, all calculations have been fully parameterized using settings recommended by the Collaborative Computational Project for NMR Crystallography (CCP-NC), as stated in an online published report.[86] The tools presented here have the advantages of harmonizing the way NMR crystallography is performed, accelerating workflow & productivity, improving the accessibility to non-specialists, and increasing the robustness of the NMR crystallography process by minimizing data manipulations by the user. The results, which can be output in Materials Studio, CASTEP, and Quantum Espresso, are also summarized in the Magres file format,[87] and can be deposited to the repository of calculated NMR parameters.[88]

Materials Studio is widely used throughout industry, and has an excellent scripting environment that allows for the development of purpose-made tools. The scripts reported here

make use of functionalities that have been made available within Materials Studio's scripting interface and are developed for the specific purpose of NMR crystallography. Although these scripts were developed with  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^{19}\text{F}$  as the target nuclei, we note that they can be adapted for use with all nuclei. Through this work, our goals are to address the current gaps faced in adopting NMR crystallography, to provide users with a starting point for coding adapted scripts, and to stimulate the sharing of NMR crystallography tools within the solid-state NMR community.

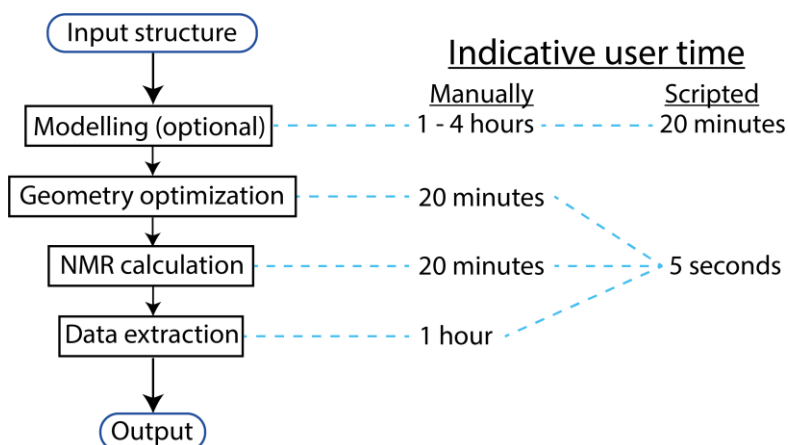
NMR crystallography makes use of experimental data, and in some of the scripts discussed herein, the user has the opportunity to input experimental chemical shifts. While this work aims to automate the typical steps involved in NMR crystallography calculations, readers interested in the experimental aspects of solid-state NMR are directed towards several resources on the topic.[89-92]

## **2 – The NMR Crystallography Toolbox.**

The typical steps involved in NMR crystallography calculations (see Figure 1) requires an input structure obtained from a diffraction experiment or crystal structure prediction (CSP).[49, 52, 62-64] As shown in Figure 1, the typical steps involved in NMR crystallography calculations start with creating a model. Usually, the user can use this input structure as is, but at times additional modelling is necessary (see section 2.3 below) and, where necessary, we've estimated that this process can take up to 4 hours of user time, e.g. modifying bond lengths, bond angles, torsion angles, or creating models of isolated molecules. We note that there are a few stumbling blocks which may have to be overcome when using a diffraction X-ray structure, namely adding any missing atoms, and removing duplicate atoms, the latter often being present when modelling crystallographic disorder. The next step in the NMR crystallography process

is geometry optimization, which allows the structure to be energetically relaxed, optimizing atomic positions, both hydrogen and heavy atom positions, and hence bond lengths. We've estimated that this step requires approximately 20 minutes of user time to fully parameterize the calculation, manually upload the job to a cluster, submit the job, and download and verify the results. Note that the user time estimates stated in this section do not include the runtime of the DFT calculations. The optimized structure is then used as the input for calculating the NMR observables, and we estimate that this also requires 20 minutes of user time to appropriately parameterize and submit the calculation. The final step of the NMR crystallography process is to extract the results and prepare the data for interpretation, which we estimate could require 1 hour of user time to complete. With the possibility of automating the process, a series of computational tools have been created and are described here, resulting in a potentially significant time saving. A summary of the scripts can be found in Table 1, while the scripts, video tutorials, and documentation on using the scripts can be downloaded (under a CC-BY 4.0 licence) from the Warwick Research Archive Portal repository at <https://wrap.warwick.ac.uk/156679>. In addition, the full text version of each script annotated with comments is available in the Supporting Information. To ensure clarity of the discussion, we have separated these tools into their respective categories: DFT calculations, data extraction & visualization, and structural modelling. We note that these tools require Materials Studio, CASTEP,[29] and ideally a server hosting the Materials Studio gateway, while the Magres2TopSpin and CSV2TopSpin scripts require the proprietary Bruker software TopSpin to function. Each script and its "user menu" dialogue is shown in the Supporting Information, and their functionalities could in principle be implemented using alternative code written in Python or Quantum Espresso.[30-32]





**Figure 1.** (left) Typical steps involved in NMR crystallography calculations. The workflow begins with an input structure (e.g. X-ray crystal structure), modelling the structure (optional step), optimizing the model, performing the NMR calculation, and extracting the data. (right) Estimated user time required to perform each step manually, and the time required when using the scripts discussed here. The user time estimates do not include the runtime of the DFT calculations.

**Table 1.** Section, title, and description of the scripts reported here.

Section	Title	Description of functionality
2.1 DFT calculations	Geometry optimization	Performs a geometry optimization of the structure using CASTEP
2.1 DFT calculations	NMR calculation	Calculates the NMR chemical shieldings & electric field gradients using CASTEP
2.1 DFT calculations	Geometry optimization & NMR calculation	Optimizes the structure and subsequently calculates the NMR chemical shieldings & electric field gradients using CASTEP
2.2 Data extraction & visualization	Extract NMR parameters	Extracts all the NMR parameters into a table with the original atom labels
2.2 Data extraction & visualization	Colouring – change in atomic coordinates	Colours atoms based on the changes in the atomic coordinates arising from the geometry optimization
2.2 Data extraction & visualization	Colouring – NMR chemical shifts (Exp. vs Calc.)	Colours atoms based on the agreement between calculated & experimental chemical shifts
2.2 Data extraction & visualization	Magres2TopSpin	Parses data from a .magres file into an input file for TopSpin’s SOLA tool
2.2 Data extraction & visualization	CSV2TopSpin	Parses data from a .CSV file into an input file for TopSpin’s SOLA tool
2.3 Structural modelling	Create isolated molecules	Creates a model of isolated molecules for each component in the structure
2.3 Structural modelling	Iterate interatomic distances	Modifies the selected distance in incremental steps

2.3 Structural modelling	Iterate interatomic angles	Modifies the selected angle in incremental steps
2.3 Structural modelling	Iterate interatomic torsion angles	Modifies the selected torsion angle in incremental steps

## 2.1 - DFT calculations

The DFT-calculations[93-95] are performed using the GIPAW[26] method as implemented in CASTEP[29] as part of Materials Studio version 17.[96] As the computational capabilities are continuously evolving, we acknowledge that flexibility has to be built into these scripts to allow for their modifications. In order to make these future-proof, the scripts have been organized so that all computational parameters, whilst defaulting to the CCP-NC recommended values,[86] can be easily modified, either through the Materials Studio interface or through the scripting environment. The GGA PBE functional[97] is employed for all calculations. The calculations are performed with TS DFT-D correction,[98] on-the-fly ultrasoft pseudopotentials, and Koelling-Harmon relativistic treatment, with the default cut-off energies,  $k$ -point separations, and optimization conditions summarized in Table 2. In addition to the calculation outputs, the DFT scripts also output a report file that summaries all the computational parameters and the optimization conditions. These computational parameters are also summarized in the .magres output file[87] which can be submitted to the CCP-NC database of calculated NMR data.[88] A series of video tutorials on performing DFT calculations using these scripts can be found in the Warwick Research Archive Portal repository, see: <https://wrap.warwick.ac.uk/156679>.

An important consideration in NMR crystallography is how to link experimental NMR chemical shifts with calculated shieldings.[34, 36, 41, 51, 99-102] Specifically, the calculated isotropic chemical shift ( $\delta_{\text{iso,calc}}$ ) can be determined from the raw calculated isotropic chemical shielding ( $\sigma_{\text{iso,calc}}$ ) using a chemical shielding reference ( $\sigma_{\text{ref}}$ ):

$$\delta_{iso,calc} = \sigma_{ref} - m\sigma_{iso,calc} \quad \text{Eq 1}$$

The scripts in this paper set the gradient  $m$  to unity. There is debate in the literature as to whether best practice is to determine  $\sigma_{ref}$  either by directly comparing the experimental and GIPAW-calculated chemical shifts for each specific compound, i.e., to determine a distinct  $\sigma_{ref}$  for each compound,[34, 36, 51, 99] or to use a single  $\sigma_{ref}$  as determined from a set of reference compounds.[41, 100-102]

**Table 2.** Default computational parameters<sup>a</sup> used in the CASTEP geometry optimization, NMR calculation, and the combination of geometry optimization followed by an NMR calculation.

parameter	geometry optimization	NMR calculation	geometry optimization & NMR calculation
cut-off energy	600 eV	700 eV	600 eV & 700 eV <sup>b</sup>
k-point separation	0.05	0.05	0.05
optimize cell	false		false
only optimize protons	false		false
use DFT-D <sup>c</sup>	true	true	true

<sup>a</sup> The settings have been fully parameterized using CCP-NC recommended settings.[86] Default parameters can be modified through the Materials Studio interface.

<sup>b</sup> The geometry optimization uses a cut-off energy of 600 eV, followed by the NMR calculation at 700 eV.

<sup>c</sup> Using the Tkatchenko-Scheffler (TS) correction.[98]

### 2.1.1 - Geometry optimization

The first step in the NMR crystallography process is to perform the geometry optimization. This task minimizes the structural energy of the system by optimizing atomic positions,[95, 103] notably those of hydrogen atoms.[77] One consideration is whether to use a fixed unit cell or to allow the unit cell dimensions to be optimized using dispersion corrected DFT.[104] The script titled “CASTEP – Geometry Optimization” begins by prompting the user with the option of modifying the cut-off energy, the  $k$ -point separation, and the usage of DFT-D employing the Tkatchenko-Scheffler (TS) [98] correction, with the default settings set

to those recommended by the CCP-NC.[86] An additional functionality allows the optimization of only the proton positions, as compared to optimizing all atoms. The script functions by parameterizing the CASTEP geometry optimization and uses the Materials Studio interface to submit the job to the defined gateway. Upon the completion of the geometry optimization, the script will output a file named “filename-Original” for the original geometry, and “filename-Optimized” containing the optimized structure. These files will serve as input files for a script described later (*vide infra*) which colours atoms based on the changes in their Cartesian atomic coordinates incurred during the geometry optimization, in order to aid in visualizing the impact of the geometry optimization. The script ends by generating a report which summarizes key computational parameters and the optimization conditions.

### 2.1.2 - NMR calculation

Upon completion of the geometry optimization, the structural model is now ready for the NMR observables to be calculated. The script titled “CASTEP - NMR Calculation” prompts the user with the option of modifying the *k*-point separation, cut-off energy, and whether to use TS DFT-D in the calculation.[98, 104] The user can simply click “OK” to accept the default settings as recommended by the CCP-NC.[86] The script functions by parameterizing a CASTEP energy calculation and uses the Materials Studio interface to submit the job to the defined gateway. A report is generated which summarizes selected computational parameters. Note that while the focus of this paper is on chemical shifts, the script also enables the calculation of the electric field gradient (EFG) which yields quadrupolar parameters for spin  $I \geq 1$  nuclei.

### 2.1.3 – Combined geometry optimization & NMR calculation

As an alternative to the user submitting the geometry optimization and then manually submitting the NMR calculation, the script titled “CASTEP – GeomOpt + NMR” automatically performs both steps sequentially, as per the discussion in sections 2.1.1 and 2.1.2. The user is prompted with a dialogue to modify the cut-off energy for the geometry optimization and NMR calculation, the *k*-point separation, and select conditions for the geometry optimization (optimize unit cell, optimize proton positions, use DFT-D). Upon completion of the geometry optimization, the optimized structure is then automatically submitted for the NMR calculation. This script includes the data extraction tool, as described below in section 2.2.1, and the results from the NMR calculation are automatically summarized in a table. Notably, this table includes the original atom labels from the original input structure (*vide infra*), thus facilitating the interpretation of the NMR results. Using the chemical shielding references provided by the user in the script’s initial dialogue, the calculated chemical shielding values are converted to chemical shifts using Eq 1. Overall, this script automates the process of NMR crystallography calculations and data extraction into a single mouse click, thus providing the user with a significant time saving.

## 2.2 – Data extraction & visualization

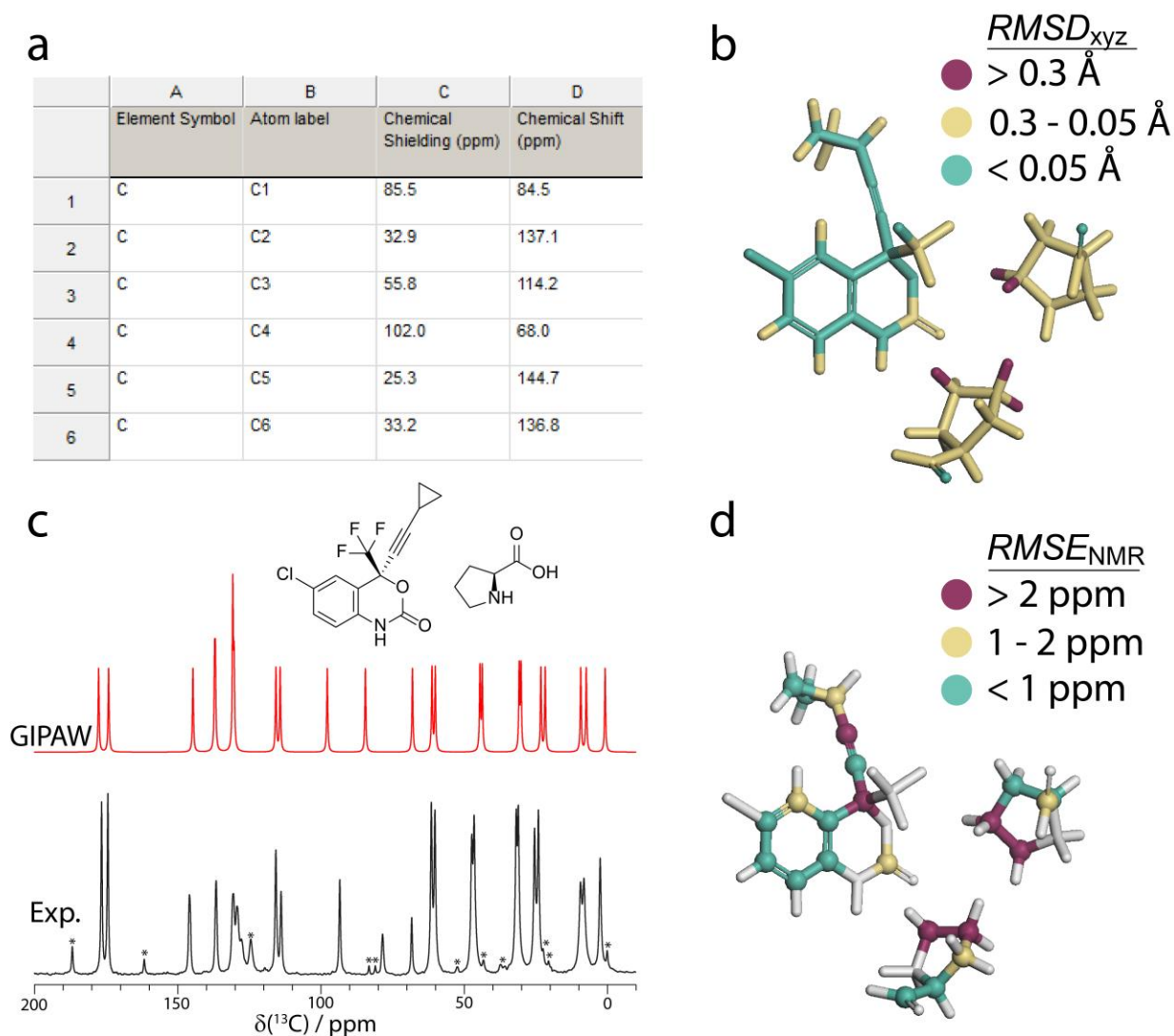
In order to complete the process of NMR crystallography, the data needs to be extracted and interpreted. The results from the CASTEP calculations are summarized in a .magres file[87] which can be easily searched using a text-based editor, or using a parsing tool. There are several tools currently available that accomplish this task, such as MagresView & MagresPython,[87] EFGshield,[105] and VMS.[106, 107] The tools that are described in this section aim to simplify the process of extracting the NMR results, interpreting the data, and communicating the results, all within Materials Studio.

### 2.2.1 – Data extraction

The NMR chemical shielding and quadrupolar coupling parameters can be quickly extracted into a table for each atom in the structure with their original atom labels using the script titled “Extract NMR parameters”. When using the scripting interface, the NMR parameters obtained from the CASTEP calculation are automatically assigned to the correct atoms in the input structure, thereby facilitating the task of creating a table with the NMR parameters and the original atom labels. The original atom labels are retained using functionalities within Materials Studio. We note that, when running stand-alone, CASTEP drops the original atom labels while generating the cell file and numbers atoms from 1 to  $X_n$  for each element, whereas crystallographic labels often uses sequential numbering for each neighboring atom regardless of whether it is a different element. If the user manually submits the CASTEP calculations, the “CASTEP analysis” tool within Materials Studio must be used prior to running the data extraction script. The analysis process assigns the calculated NMR parameters from the CASTEP output file to the atoms in the input structure, which allows the script to access the NMR parameters from the atom properties.

The “Extract NMR parameters” script functions by first creating an array which includes all atoms in the structure. The script then surveys each of the atoms in the array individually and extracts the element symbol, the atom name, the isotropic chemical shielding, the quadrupolar coupling constant ( $C_Q$ ), and the quadrupolar asymmetry parameter ( $\eta$ ). The script also converts the chemical shielding values into chemical shifts using the user-defined chemical shielding references, and summarizes all the data in a table, as shown in Figure 2a. The script outputs the chemical shifts for  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{17}\text{O}$ , and  $^{19}\text{F}$ , where these elements are present in the system. If additional nuclei are required, the user can modify the script to include

their desired nuclei or calculate the chemical shifts manually. This process is similar to the functionality implemented in MagresView,[87] where chemical shifts can be visualized on the 3D structure, and exported into a table.



**Figure 2.** Result of the data extraction & visualization scripts. (a) Table containing the element symbol, atomic labels, NMR chemical shielding, and NMR chemical shifts ( $\sigma_{\text{ref}} = 170$  ppm for  $^{13}\text{C}$ ) obtained from the “Extract NMR Parameters” script (section 2.2.1). (b) Asymmetric unit of the cocrystal (Efavirenz)(L-proline),[108] colouring the atoms based on the  $RMSD_{xyz}$  in the Cartesian coordinates between the optimized and unoptimized structure using the “Colouring\_Atomic Coordinates” (section 2.2.2). (c) Experimental and simulated (for the GIPAW-calculated chemical shifts)  $^{13}\text{C}$  solid-state NMR cross-polarization magic-angle spinning spectrum ( $\nu_L(^1\text{H}) = 500$  MHz,  $\nu_{\text{MAS}} = 11750$  Hz, contact time = 2 ms, asterisks denote spinning sidebands that were subsequently added during the figure preparation) of the cocrystal (Efavirenz)(L-proline) using the Magres2TopSpin script (section 2.2.4). (d) Asymmetric unit of (Efavirenz)(L-proline) with the carbon atoms coloured based on the  $RMSE_{\text{NMR}}$  using the “NMR\_Colouring” scripts (section 2.2.3). The experimental and GIPAW-calculated NMR data is taken from Szell et al.[109]

### 2.2.2 – Colour atoms based on the changes in their atomic coordinates

As also discussed in Section 2.1.1, the geometry optimization process will cause changes to the atomic coordinates,[45, 110-114] and a tool that highlights these changes can bring attention to areas having incurred significant changes. In addition, geometry optimization can at times cause large unexpected changes in the structure to occur, and a method for quickly visualizing the magnitude of the change in the atomic coordinates can be helpful in easily identifying potential errors in the input structure. To this end, the script titled “Colouring\_Atomic Coordinates” has been written to colour atoms based on the differences in their Cartesian coordinates between the starting point and the optimized structure. The script functions by creating an array of all atoms in the optimized structure and the unoptimized structure. The script then matches the atoms from the optimized and unoptimized structures, and calculates the differences in their  $x$ ,  $y$ , and  $z$  Cartesian coordinates. The root mean square deviation in the Cartesian coordinates ( $RMSD_{xyz}$ ) is calculated in Eq 2 using these values as input, and a colour is assigned based on the  $RMSD_{xyz}$ , as shown in Figure 2b.

$$RMSD_{xyz} = \sqrt{\frac{1}{3}[(x_{new} - x_{old})^2 + (y_{new} - y_{old})^2 + (z_{new} - z_{old})^2]} \quad \text{Eq 2}$$

In Eq 2, the  $RMSD_{xyz}$  is calculated using the Cartesian coordinates of the refined structure ( $x_{new}$ ,  $y_{new}$ , and  $z_{new}$ ) and the original structure ( $x_{old}$ ,  $y_{old}$ , and  $z_{old}$ ). As set by default, if the  $RMSD_{xyz}$  is lower than 0.05 Å, signifying only a small change in the Cartesian coordinates, the atom is coloured green. If the  $RMSD_{xyz}$  is greater than 0.3 Å, the atom is coloured red. If the  $RMSD_{xyz}$  is greater than 0.05 Å but lower than 0.3 Å, the atom is coloured yellow. This simple visualization aid thus highlights large changes in the Cartesian coordinates that result from the geometry optimization, allowing the user to easily identify potential errors in the input structure. As shown in Figure 2b, the atoms in the cocrystal (Efavirenz)(*L*-proline) (Cambridge Structural Database[115] reference HUDRAC[108]) are coloured based on the differences in



their Cartesian coordinates between the CSD structure and the optimized structure.[109] The largest changes in the Cartesian coordinates can be seen for the hydrogen positions of the *L*-proline molecules, as shown in red, and smaller changes can be seen for the carbon atoms and hydrogen positions of the *L*-proline molecules, as shown in yellow.

### *2.2.3 – Colour atoms based on the agreement between their experimental & calculated chemical shifts*

Communicating NMR crystallography results to the non-NMR specialist is a challenge that can be facilitated through the use of visualization tools. For instance, an ORTEP-like plot can be created as part of the NMR crystallography process to show the uncertainty on the atomic positions.[68] Here, a two-step process has been created to visualize the agreement between the experimental & computational chemical shifts using the same red-yellow-green atom colouring strategy as discussed in section 2.2.2, titled “NMR\_Colouring\_Step1\_Data-Extraction” and “NMR\_Colouring\_Step2\_Data-Workup”. In the first step, the user is requested to select an element and enter the chemical shielding reference. The first script extracts the NMR data for each atom of the desired element and creates a table summarizing their calculated chemical shifts. For the current functionality, this requires the user to manually input the experimental chemical shifts. The script uses a similar method as the data extraction script from section 2.2.1, and creates an array which includes all atoms in the structure. The script creates a table, filters for the atoms corresponding to the desired element, extracts the isotropic chemical shielding values, calculates the chemical shift using the chemical shielding reference, and adds the data to the table. The user must then start the second script which calculates the absolute difference between the experimental and calculated NMR chemical shift for each atom individually, and colours the atoms based on these differences. The user can

specify the limits for colouring the atoms, which can vary from nucleus to nucleus. As set by default, if the difference between experimental and calculated chemical shift is lower than 1 ppm, the atom is coloured green. If the difference is greater than 2 ppm, the atom is coloured red. If the difference is between 1 ppm and 2 ppm, a typical discrepancy for  $^{13}\text{C}$ [28, 33-38], the atom is coloured in yellow.

In addition, the total root-mean square error ( $RMSE_{\text{NMR}}$ ) between the experimental ( $\delta_{\text{exp}}$ ) and calculated chemical shifts ( $\delta_{\text{calc}}$ ) of the entire structure is determined using Eq 3, as introduced by Emsley and co-workers[116] for ranking CSP generated structures using  $^1\text{H}$  NMR, thus allowing the structure to be scored based on the agreement between the experimental and theoretical results.

$$RMSE_{\text{NMR}} = \sqrt{\frac{1}{N} \sum_i^N (\delta_{\text{calc},i} - \delta_{\text{exp},i})^2} \quad \text{Eq 3}$$

In Eq 3,  $N$  denotes the number of atoms being surveyed. As shown in Figure 2d, the atoms in the structure are coloured based on the agreement between their experimental and calculated  $^{13}\text{C}$  chemical shifts. In the example of the (Efavirenz)(*L*-proline) cocrystal,[108, 109] the atoms with large differences between experimental and calculated chemical shifts are shown in red, atoms in moderate agreement are shown in yellow, and atoms in good agreement are shown in green. The grey atoms denote non-carbon atoms, and the  $\text{CF}_3$  carbon which was excluded due to uncertainty on the chemical shift arising from  $^{19}\text{F}$ - $^{13}\text{C}$   $J$ -coupling and spectral overlap. Using this visualization approach, NMR crystallography results can be quickly communicated in a way that can be understood by the non-specialist. We note that approaches to rank crystal structures based on experimental and calculated chemical shifts have been explored in the literature,[37, 49, 50, 52, 64, 116, 117] and the scripts shown here can be modified by the user to implement additional analysis.

#### 2.2.4 – *Magres2TopSpin* & *CSV2TopSpin*

A second method of communicating the agreement between experimental and calculated NMR results is to overlay the experimental NMR spectrum with a simulated spectrum generated from the calculated chemical shifts. One approach to accomplish this task is to add a “sticks” representation of the calculated chemical shifts underneath the experimental spectrum, as implemented in MagresView.[87] However, this is a process that requires several steps, and cannot be manipulated directly within the proprietary Bruker software, *TopSpin*, used for recording and processing NMR data. Consequently, two scripts were created to facilitate the simulation of the NMR calculated results directly within TopSpin. Once the user has added the scripts to TopSpin’s Python folder, they can be called by entering the commands *Magres2Topspin* or *CSV2Topspin*. The user is prompted with several parameters (see Figures S9 and S10 of the Supporting Information), including the location and name of the .magres or .CSV file, the name of the output model, and the amount of line broadening to be used in the simulation. In the case of the *Magres2TopSpin* script, the user is also required to enter the element type and the  $\sigma_{\text{ref}}$  in order to convert the  $\sigma_{\text{iso}}$  into  $\delta_{\text{iso}}$  using Eq 1. In the case of the *CSV2TopSpin* script, the user simply has to create a list of  $\delta_{\text{iso}}$  in a CSV file that is delimited by carriage returns. The purpose of the *CSV2TopSpin* script is to give the user the ability to simulate a list of chemical shifts, such as when using computational tools that don’t use the .magres format. After using the *Magres2TopSpin* or *CSV2TopSpin* script, the user must then open the Solids Lineshape Analysis (SOLA) tool as integrated into Bruker’s TopSpin[85] and read the input files that were created by the script. To get around SOLA’s limit of 32 spins, the scripts generate several input files containing sets of 32 spins, which can be co-added to produce the final simulated spectrum. Both scripts support input files with up to 512 spins.

The advantage of this approach is to centralize the workflow within TopSpin, allowing data processing to be expedited. The Magres2TopSpin script functions by parsing data from the defined .magres file, which contains a summary of the calculated NMR data, into an input file that can be read by the SOLA tool within TopSpin. The CSV2TopSpin script works using a similar approach, but will instead convert the chemical shifts from a .CSV file into an input file that can be read by SOLA. Further, SOLA has built-in capabilities of flexible parameters that can be modified by the user, such as the line broadening and peak intensities, and can simulate MAS sideband intensities. As shown in Figure 2c, an NMR spectrum for the (Efavirenz)(L-proline)[108, 109] cocrystal was simulated from the .magres file, with the process taking only a few seconds.

### *2.3 - Structural Modelling*

For more advanced NMR crystallography requirements, modifying the structural model may be necessary. This may be due to the presence of specific areas of uncertainty in the structure, such as the proton positions or the occurrence of crystallographic disorder or errors in the crystal structure. In order to alleviate the task of creating the structural models, a series of scripts was created to streamline the process of structural modelling. These are discussed in the following section.

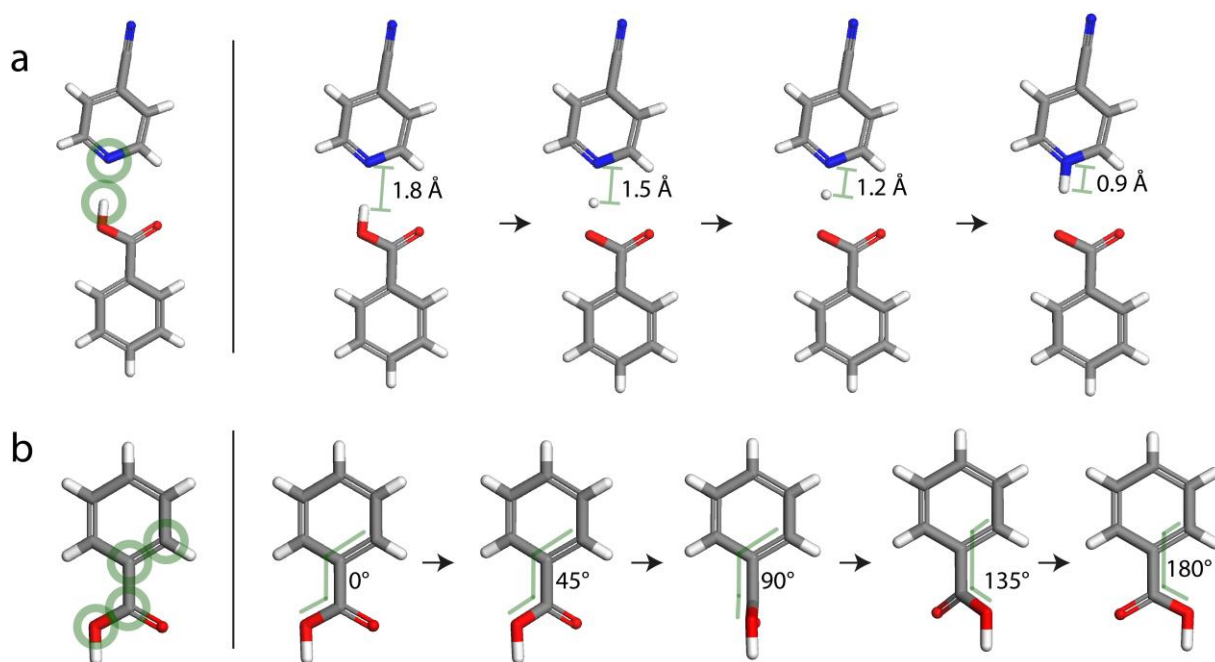
#### *2.3.1 – Iterative geometrical changes*

Sampling the conformational landscape can be useful for interpreting the influence of bond lengths, bond angles, and torsion angles on the NMR observables and the total energy.[68, 109, 118-122] For instance, knowing the change in the  $^{15}\text{N}$  or  $^1\text{H}$  chemical shift as a function of the  $\text{N}\cdots\text{H}$  distance in the  $\text{N}\cdots\text{H}\cdots\text{Y}$  motif ( $\text{Y}$  = substituent) can be an effective approach to investigate whether a particular compound is a salt or a cocrystal. Alternatively, this approach

can be used to sample the total structural energy profile as a function of a geometrical parameter (bond distance, bond angle, torsion angle), allowing energy barriers and kinetics to be estimated. However, this process can be time consuming, and a series of scripts have been created to automate this process.

Three scripts have been created, allowing the user to make changes to an interatomic distance, angle, or torsion angle. Each script begins by prompting the user to enter the name of the distance/angle/torsion that is to be varied, and to define the starting value, the ending value, and the step size. Additional options include adding a constraint to the geometrical parameter, and rebuilding the cell to create a fully periodic structural model ready for the CASTEP calculations. Upon pressing the “OK” button, the script will output a series of structures following the user parameters, and these structures can be used as input files for the CASTEP calculations. The script functions by iteratively modifying the geometrical parameter (distance, angle, torsion angle) applying the user defined parameters, adds the geometrical constraints (if requested), and creates a new document containing the geometry of each iteration. Each new document has a unique name that reflects the geometrical parameter and its value, and also eliminates the risk of overwriting the original file.

As shown in Figure 3a, the script was applied to the H···N atoms in a cocrystal of (benzoic acid)(4-cyanopyridine),[123] modifying the distance from 1.8 Å (cocrystal form) to 1.0 Å (salt form) in steps of 0.3 Å. In this example, the intended use of these models is to sample the calculated <sup>15</sup>N chemical shift in the continuum between the cocrystal and the salt form. A second example is shown in Figure 3b, where the torsion angle between the carboxylic acid group and the benzene ring in benzoic acid[124] is modified in steps of 45°, from 0° to 180°, causing the carboxylic acid to rotate through 180°. This approach can be useful in sampling the conformational landscape of rotating moieties, such as CH<sub>3</sub>, CF<sub>3</sub>, SO<sub>3</sub>, and phenyl groups.

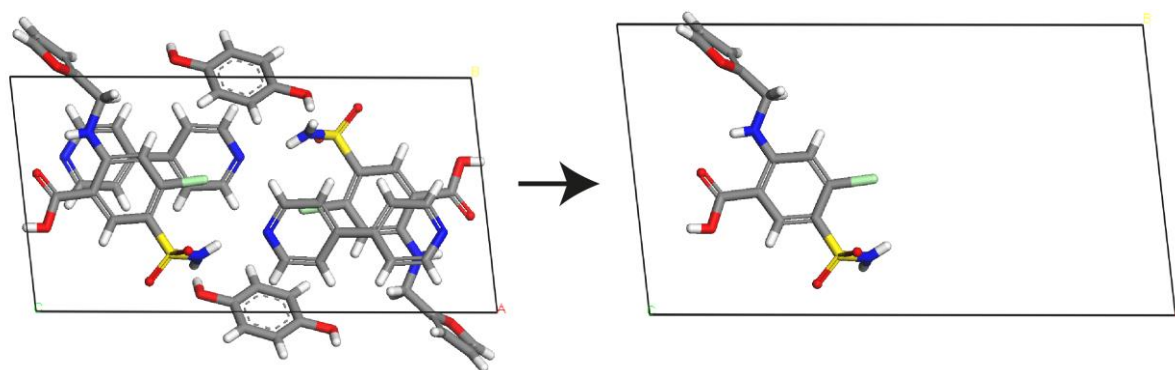


**Figure 3.** (a) Implementation of the “Set distances” script as described in section 2.3.1 on the  $N\cdots H$  distance in structure PUKPER,[123] a cocrystal of (benzoic acid)(4-cyanopyridine), from the cocrystal form ( $d_{N\cdots H} = 1.8 \text{ \AA}$ ) to the salt form ( $d_{N\cdots H} = 1.0 \text{ \AA}$ ). (b) Implementation of the “Set torsions” script as described in section 2.3.1, rotating the carboxylic acid group of structure BENZAC20[124] by modifying the C-C-C-O torsion angle.

### 2.3.2 – Isolated molecules

Creating models of isolated molecules can be helpful, for instance, in finding fingerprints of intermolecular interactions, notably ring currents, and hydrogen bonds using NMR-calculated chemical shifts.[38, 51, 125-132] To create models of isolated molecules, the space group is generally modified to *P1* in order to eliminate symmetry-related molecules from the unit cell, and the unit cell size is usually increased to reduce the occurrence of through-space interactions between unit cells. Due to the rising computational costs associated with increasing the size of the unit cell, it is undesirable to increase the cell size by more than 10 Å in all directions, as has been shown by Tatton et al.[86] The script prompts the user to input the length to be added to the *a*, *b*, and *c* axes of the unit cell, whether to create the new structures in the *P1* space group, and whether to isolate each molecule individually or to isolate the entire

unit cell. The script functions by first unbuilding the crystal into a non-periodic structure, and then modifies the  $a$ ,  $b$ , and  $c$  axes of the unit cell parameters to their desired values. The script then creates an array containing the unique molecules from the structure, and outputs each molecule individually into a new file. If it has been specified by the user, the unit cell will be rebuilt using the new unit cell parameters and the  $P1$  space group. As shown in Figure 4, the script titled “Create isolated molecules” allows the user to quickly create models of isolated molecules, demonstrating the script on a crystal structure of a ternary furosemide cocrystal (CSD[115] reference BOKHUG[133]). In this case, the script creates a model for each of the three unique molecules in the asymmetric unit.



**Figure 4.** (left) Unmodified unit cell of a ternary furosemide cocrystal (BOKHUG[133]), and (right) isolated molecule of furosemide in an enlarged unit cell of the same structure obtained from the “Create isolated molecule” script as described in section 2.3.2. The isolated molecule is in the  $P1$  space group.

#### 2.4 - Additional resources & tutorials

NMR crystallography can be challenging. Although we hope that these scripts will facilitate the process of NMR crystallography, we acknowledge that some expert knowledge of NMR crystallography is still required. As a result, we have created a series of video tutorials dedicated to training users and have fully documented how to install and modify these tools. The documentation and video tutorials can be found in the Warwick Research Archive Portal repository (see: <https://wrap.warwick.ac.uk/156679>). Additional details on the scripts,

including the full text version of each script with comments, are available in the Supporting Information, and users are encouraged to modify these tools to add their desired functionalities. These scripts have been created and shared on a free-to-use basis, and made available at WRAP under a CC BY 4.0 licence.[134] The authors and owners of the scripts permit the modification of the scripts on the condition that appropriate credit is given to the authors.

## **Conclusions**

A series of computational tools has been created with the purpose of addressing some of the bottlenecks associated with carrying out NMR crystallography. These tools address: CASTEP job submission, data extraction & visualization, and modelling. In the Materials Studio environment, we show that the NMR crystallography process can be performed in as little as a single mouse click, with a single script submitting the structure for an optimization and then automatically submitting the optimized structure for the NMR calculation. Further, we have scripted several useful tools to expedite the process of NMR crystallography, such as extracting the NMR results for each atom in the structure, visualizing the results using a traffic-light system, and modelling the structure. With the growing interest and capabilities of NMR crystallography, we hope that this work will help expedite the process for the user and stimulate the sharing of purpose-made tools. For instance, the automation of more advanced NMR crystallography schemes through scripting could benefit users with a higher accessibility and ease of use, while reducing the likelihood of errors. The community is encouraged to share their new tools through the CCP-NC (see: <https://www.ccpnc.ac.uk/>).

**Supplementary Information.** The supplementary information contains supporting documentation, training videos, and scripts.



**Acknowledgements.** This work was funded by Innovate UK and AstraZeneca (Grant number: KTP11570). The authors thank Matthew Pacitto for helpful discussions for the Magres2TopSpin parsing script. Data for this study are provided as a supporting data set from WRAP, the Warwick Research Archive Portal at <https://wrap.warwick.ac.uk/156679>.

**Author Contributions.** The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

**Conflicts of Interest.** The authors declare no conflicts of interest.

## References

1. F. Atassi, C. Mao, A.S. Masadeh, S.R. Byrn, *Solid-state characterization of amorphous and mesomorphous calcium ketoprofen*, *J. Pharm. Sci.* 99 (2010) 3684-3697.
2. A. Berzins, P. Hodgkinson, *Solid-state NMR and computational investigation of solvent molecule arrangement and dynamics in isostructural solvates of droperidol*, *Solid State Nucl. Magn. Reson.* 65 (2015) 12-20.
3. D.C. Apperley, P.A. Basford, C.I. Dallman, R.K. Harris, M. Kinns, P.V. Marshall, A.G. Swanson, *Nuclear magnetic resonance investigation of the interaction of water vapor with sildenafil citrate in the solid state*, *J. Pharm. Sci.* 94 (2005) 516-523.
4. D.E. Braun, L.H. Koztecki, J.A. McMahon, S.L. Price, S.M. Reutzel-Edens, *Navigating the Waters of Unconventional Crystalline Hydrates*, *Mol. Pharmaceutics* 12 (2015) 3069-3088.
5. S.D. Gumbert, M. Körbitzer, E. Alig, M.U. Schmidt, M.R. Chierotti, R. Gobetto, X. Li, J. van de Streek, *Crystal structure and tautomerism of Pigment Yellow 138 determined by X-ray powder diffraction and solid-state NMR*, *Dyes Pigm.* 131 (2016) 364-372.
6. P. Cerreia Vioglio, M.R. Chierotti, R. Gobetto, *Pharmaceutical aspects of salt and cocrystal forms of APIs and characterization challenges*, *Adv. Drug Deliv. Rev.* 117 (2017) 86-110.
7. B. Elena, G. Pintacuda, N. Mifsud, L. Emsley, *Molecular structure determination in powders by NMR crystallography from proton spin diffusion*, *J. Am. Chem. Soc.* 128 (2006) 9555-9560.
8. R.K. Harris, *Applications of solid-state NMR to pharmaceutical polymorphism and related matters*, *J. Pharm. Pharmacol.* 59 (2007) 225-239.
9. K.M. Burgess, F.A. Perras, A. Lebrun, E. Messner-Henning, I. Korobkov, D.L. Bryce, *Sodium-23 solid-state nuclear magnetic resonance of commercial sodium naproxen and its solvates*, *Journal of Pharmaceutical Sciences* 101 (2012) 2930-2940.
10. P.M.J. Szell, S. Zablony, D.L. Bryce, *Halogen bonding as a supramolecular dynamics catalyst*, *Nat. Commun.* 10 (2019) 916.

11. D.L.M. de Aguiar, R.A. da Silva San Gil, R.B. de Alencastro, E.F. de Souza, B.L. Borré, V. da Silva Vaiss, A.A. Leitão, *6-Aminopenicillanic acid revisited: A combined solid state NMR and in silico refinement*, Chem. Phys. Lett. 660 (2016) 214-220.
12. M. Zilka, J.R. Yates, S.P. Brown, *An NMR crystallography investigation of furosemide*, Magn. Reson. Chem. 57 (2019) 191-199.
13. N.J. Vigilante, M.A. Mehta, *A <sup>13</sup>C solid-state NMR investigation of four cocrystals of caffeine and theophylline*, Acta Cryst. C73 (2017) 234-243.
14. S.P. Brown, *Applications of high-resolution <sup>1</sup>H solid-state NMR*, Solid State Nucl. Magn. Reson. 41 (2012) 1-27.
15. J. Brus, J. Czernek, M. Hruby, P. Svec, L. Kobera, S. Abbrent, M. Urbanova, *Efficient Strategy for Determining the Atomic-Resolution Structure of Micro- and Nanocrystalline Solids within Polymeric Microbeads: Domain-Edited NMR Crystallography*, Macromolecules 51 (2018) 5364-5374.
16. D.L. Bryce, *New frontiers for solid-state NMR across the periodic table: a snapshot of modern techniques and instrumentation*, Dalton Trans. 48 (2019) 8014-8020.
17. P.A. Tishmack, D.E. Bugay, S.R. Byrn, *Solid-state nuclear magnetic resonance spectroscopy--pharmaceutical applications*, J. Pharm. Sci. 92 (2003) 441-474.
18. D.H. Barich, J.M. Davis, L.J. Schieber, M.T. Zell, E.J. Munson, *Investigation of solid-state NMR line widths of ibuprofen in drug formulations*, J. Pharm. Sci. 95 (2006) 1586-1594.
19. M. Geppi, G. Mollica, S. Borsacchi, C.A. Veracini, *Solid-State NMR Studies of Pharmaceutical Systems*, Appl. Spectrosc. Rev. 43 (2008) 202-302.
20. F.G. Vogt, *Evolution of solid-state NMR in pharmaceutical analysis*, Future Med. Chem. 2 (2010) 915-921.
21. R.K. Harris, S.A. Joyce, C.J. Pickard, S. Cadars, L. Emsley, *Assigning carbon-13 NMR spectra to crystal structures by the INADEQUATE pulse sequence and first principles computation: a case study of two forms of testosterone*, Phys. Chem. Chem. Phys. 8 (2006) 137-143.
22. E. Carignani, S. Borsacchi, M. Geppi, *Dynamics by solid-state NMR: detailed study of ibuprofen Na salt and comparison with ibuprofen*, J. Phys. Chem. A 115 (2011) 8783-8790.
23. L. Zhao, A.C. Pinon, L. Emsley, A.J. Rossini, *DNP-enhanced solid-state NMR spectroscopy of active pharmaceutical ingredients*, Magn. Reson. Chem. 56 (2018) 583-609.
24. R.K. Harris, P. Hodgkinson, V. Zorin, J.N. Dumez, B. Elena-Herrmann, L. Emsley, E. Salager, R.S. Stein, *Computation and NMR crystallography of terbutaline sulfate*, Magn. Reson. Chem. 48 Suppl 1 (2010) S103-112.
25. X. Lu, Y. Tsutsumi, C. Huang, W. Xu, S.R. Byrn, A.C. Templeton, A.V. Buevich, J.P. Amoureux, Y. Su, *Molecular packing of pharmaceuticals analyzed with paramagnetic relaxation enhancement and ultrafast magic angle pinning NMR*, Phys. Chem. Chem. Phys. 22 (2020) 13160-13170.
26. C.J. Pickard, F. Mauri, *All-electron magnetic response with pseudopotentials: NMR chemical shifts*, Phys. Rev. B 63 (2001) 245101.
27. T. Charpentier, *The PAW/GIPAW approach for computing NMR parameters: a new dimension added to NMR study of solids*, Solid State Nucl. Magn. Reson. 40 (2011) 1-20.
28. C. Bonhomme, C. Gervais, F. Babonneau, C. Coelho, F. Pourpoint, T. Azais, S.E. Ashbrook, J.M. Griffin, J.R. Yates, F. Mauri, et al., *First-principles calculation of NMR parameters using the gauge including projector augmented wave method: a chemist's point of view*, Chem. Rev. 112 (2012) 5733-5779.
29. S.J. Clark, M.D. Segall, C.J. Pickard, P.J. Hasnip, M.J. Probert, K. Refson, M.C. Payne, *First principles methods using CASTEP*, Z. Kristallogr. 220 (2005) 567-570.
30. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, et al., *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*, J. Phys.: Condens. Matter 21 (2009) 395502.

31. P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, et al., *Advanced capabilities for materials modelling with Quantum ESPRESSO*, J. Phys.: Condens. Matter 29 (2017) 465901.
32. P. Giannozzi, O. Baseggio, P. Bonfa, D. Brunato, R. Car, I. Carnimeo, C. Cavazzoni, S. de Gironcoli, P. Delugas, F. Ferrari Ruffino, et al., *Quantum ESPRESSO toward the exascale*, J. Chem. Phys. 152 (2020) 154105.
33. E.K. Corlett, H. Blade, L.P. Hughes, P.J. Sidebottom, D. Walker, R.I. Walton, S.P. Brown, *Investigating discrepancies between experimental solid-state NMR and GIPAW calculation: N=C-N <sup>13</sup>C and OH...O <sup>1</sup>H chemical shifts in pyridinium fumarates and their cocrystals*, Solid State Nucl. Magn. Reson. 108 (2020) 101662.
34. R.K. Harris, P. Hodgkinson, C.J. Pickard, J.R. Yates, V. Zorin, *Chemical shift computations on a crystallographic basis: some reflections and comments*, Magn. Reson. Chem. 45 Suppl 1 (2007) S174-S186.
35. S.E. Ashbrook, P. Hodgkinson, *Perspective: Current advances in solid-state NMR spectroscopy*, J. Chem. Phys. 149 (2018) 040901.
36. P. Hodgkinson, *NMR crystallography of molecular organics*, Prog. Nucl. Magn. Reson. Spectrosc. 118-119 (2020) 10-53.
37. F.M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti, L. Emsley, *Chemical shifts in molecular solids by machine learning*, Nat. Commun. 9 (2018) 4501.
38. H. Blade, C.D. Blundell, S.P. Brown, J. Carson, H.R.W. Dannatt, L.P. Hughes, A.K. Menakath, *Conformations in Solution and in Solid-State Polymorphs: Correlating Experimental and Calculated Nuclear Magnetic Resonance Chemical Shifts for Tolfenamic Acid*, J. Phys. Chem. A 124 (2020) 8959-8977.
39. D.L. Bryce, *NMR crystallography: structure and properties of materials from solid-state nuclear magnetic resonance observables*, IUCr 4 (2017) 350-359.
40. S.E. Ashbrook, D. McKay, *Combining solid-state NMR spectroscopy with first-principles calculations - a guide to NMR crystallography*, Chem. Commun. 52 (2016) 7186-7204.
41. G.J. Beran, J.D. Hartman, Y.N. Heit, *Predicting Molecular Crystal Properties from First Principles: Finite-Temperature Thermochemistry to NMR Crystallography*, Acc. Chem. Res. 49 (2016) 2501-2508.
42. M.R. Chierotti, R. Gobetto, *NMR crystallography: the use of dipolar interactions in polymorph and co-crystal investigation*, CrystEngComm 15 (2013) 8599-8612.
43. F. Taulelle, *NMR crystallography: crystallochemical formula and space group selection*, Solid State Sci. 6 (2004) 1053-1057.
44. R.K. Harris, *NMR crystallography: the use of chemical shifts*, Solid State Sci. 6 (2004) 1025-1037.
45. E. Salager, R.S. Stein, C.J. Pickard, B. Elena, L. Emsley, *Powder NMR crystallography of thymol*, Phys. Chem. Chem. Phys. 11 (2009) 2610-2621.
46. R.K. Harris, *NMR studies of organic polymorphs and solvates*, Analyst 131 (2006) 351-373.
47. M. Schmidt, C.S. Zehe, R. Siegel, J.U. Heigl, C. Steinlein, H.-W. Schmidt, J. Senker, *NMR-crystallographic study of two-dimensionally self-assembled cyclohexane-based low-molecular-mass organic compounds*, CrystEngComm 15 (2013) 8784-8796.
48. S.A. Southern, D.L. Bryce, *Recent advances in NMR crystallography and polymorphism*, Annu. Rep. NMR Spectrosc. 102 (2021) 1-80.
49. A. Hofstetter, M. Balodis, F.M. Paruzzo, C.M. Widdifield, G. Stevanato, A.C. Pinon, P.J. Bygrave, G.M. Day, L. Emsley, *Rapid Structure Determination of Molecular Solids Using Chemical Shifts Directed by Unambiguous Prior Constraints*, J Am Chem Soc 141 (2019) 16624-16634.
50. C.J. Pickard, E. Salager, G. Pintacuda, B. Elena, L. Emsley, *Resolving structures from powders by NMR crystallography using combined proton spin diffusion and plane wave DFT calculations*, J. Am. Chem. Soc. 129 (2007) 8932-8933.

51. A.L. Webber, L. Emsley, R.M. Claramunt, S.P. Brown, *NMR crystallography of campho[2,3-c]pyrazole ( $Z' = 6$ ): combining high-resolution  $^1\text{H}$ - $^{13}\text{C}$  solid-state MAS NMR spectroscopy and GIPAW chemical-shift calculations*, *J. Phys. Chem. A* 114 (2010) 10435-10442.
52. M. Baias, C.M. Widdifield, J.N. Dumez, H.P. Thompson, T.G. Cooper, E. Salager, S. Bassil, R.S. Stein, A. Lesage, G.M. Day, et al., *Powder crystallography of pharmaceutical materials by combined crystal structure prediction and solid-state  $^1\text{H}$  NMR spectroscopy*, *Phys. Chem. Chem. Phys.* 15 (2013) 8069-8080.
53. S.O. Nilsson Lill, C.M. Widdifield, A. Pettersen, A. Svensk Ankarberg, M. Lindkvist, P. Aldred, S. Gracin, N. Shankland, K. Shankland, S. Schantz, et al., *Elucidating an Amorphous Form Stabilization Mechanism for Tenapanor Hydrochloride: Crystal Structure Analysis Using X-ray Diffraction, NMR Crystallography, and Molecular Modeling*, *Mol. Pharmaceutics* 15 (2018) 1476-1487.
54. R.K. Harris, S. Cadars, L. Emsley, J.R. Yates, C.J. Pickard, R.K. Jetti, U.J. Griesser, *NMR crystallography of oxybuprocaine hydrochloride, Modification II $^\circ$* , *Phys. Chem. Chem. Phys.* 9 (2007) 360-368.
55. A. Kumar, B.J. Walder, A. Kunhi Mohamed, A. Hofstetter, B. Srinivasan, A.J. Rossini, K. Scrivener, L. Emsley, P. Bowen, *The Atomic-Level Structure of Cementitious Calcium Silicate Hydrate*, *J. Phys. Chem. C* 121 (2017) 17188-17196.
56. A.C. Pinon, A.J. Rossini, C.M. Widdifield, D. Gajan, L. Emsley, *Polymorphs of Theophylline Characterized by DNP Enhanced Solid-State NMR*, *Mol. Pharmaceutics* 12 (2015) 4146-4153.
57. A.J. Rossini, A. Zagdoun, F. Hegner, M. Schwarzwald, D. Gajan, C. Coperet, A. Lesage, L. Emsley, *Dynamic nuclear polarization NMR spectroscopy of microcrystalline solids*, *J. Am. Chem. Soc.* 134 (2012) 16899-16908.
58. N.M. Dicaire, F.A. Perras, D.L. Bryce,  *$^{23}\text{Na}$  magic-angle spinning and double-rotation NMR study of solid forms of sodium valproate*, *Can. J. Chem.* 92 (2014) 9-15.
59. C.M. Widdifield, S.O. Nilsson Lill, A. Broo, M. Lindkvist, A. Pettersen, A. Svensk Ankarberg, P. Aldred, S. Schantz, L. Emsley, *Does  $Z'$  equal 1 or 2? Enhanced powder NMR crystallography verification of a disordered room temperature crystal structure of a p38 inhibitor for chronic obstructive pulmonary disease*, *Phys. Chem. Chem. Phys.* 19 (2017) 16650-16661.
60. P.M.J. Szell, S.P. Brown, L.P. Hughes, H. Blade, S.O. Nilsson Lill, *A curious case of dynamic disorder in pyrrolidine rings elucidated by NMR crystallography*, *Chem. Commun.* 56 (2020) 14039-14042.
61. T. Pawlak, I. Sudgen, G. Bujacz, D. Iuga, S.P. Brown, M.J. Potrzebowski, *Synergy of Solid-State NMR, Single-Crystal X-ray Diffraction, and Crystal Structure Prediction Methods: A Case Study of Teriflunomide (TFM)*, *Cryst. Growth Des.* 21 (2021) 3328-3343.
62. M. Cordova, M. Balodis, A. Hofstetter, F. Paruzzo, S.O. Nilsson Lill, E.S.E. Eriksson, P. Berruyer, B. Simoes de Almeida, M.J. Quayle, S.T. Norberg, et al., *Structure determination of an amorphous drug through large-scale NMR predictions*, *Nat. Commun.* 12 (2021) 2964.
63. J. Brus, J. Czernek, L. Kobera, M. Urbanova, S. Abbrent, M. Husak, *Predicting the Crystal Structure of Decitabine by Powder NMR Crystallography: Influence of Long-Range Molecular Packing Symmetry on NMR Parameters*, *Cryst. Growth Des.* 16 (2016) 7102-7111.
64. M. Baias, J.N. Dumez, P.H. Svensson, S. Schantz, G.M. Day, L. Emsley, *De novo determination of the crystal structure of a large drug molecule by crystal structure prediction-based powder NMR crystallography*, *J. Am. Chem. Soc.* 135 (2013) 17501-17507.
65. K. Kalakewich, R. Iulicci, K.T. Mueller, H. Eloranta, J.K. Harper, *Monitoring the refinement of crystal structures with  $^{15}\text{N}$  solid-state NMR shift tensor data*, *J. Phys. Chem.* 143 (2015) 194702.
66. C.M. Widdifield, F.A. Perras, D.L. Bryce, *Solid-state  $^{185/187}\text{Re}$  NMR and GIPAW DFT study of perhenates and  $\text{Re}_2(\text{CO})_{10}$ : chemical shift anisotropy, NMR crystallography, and a metal-metal bond*, *Phys. Chem. Chem. Phys.* 17 (2015) 10118-10134.

67. E.A. Engel, A. Anelli, A. Hofstetter, F. Paruzzo, L. Emsley, M. Ceriotti, *A Bayesian approach to NMR crystal structure determination*, *Phys. Chem. Chem. Phys.* 21 (2019) 23385-23400.
68. A. Hofstetter, L. Emsley, *Positional Variance in NMR Crystallography*, *J. Am. Chem. Soc.* 139 (2017) 2573-2576.
69. F.A. Perras, I. Korobkov, D.L. Bryce, *NMR crystallography of sodium diphosphates: combining dipolar, shielding, quadrupolar, diffraction, and computational information*, *CrystEngComm* 15 (2013) 8727-8738.
70. A.S. Tatton, T.N. Pham, F.G. Vogt, D. Iuga, A.J. Edwards, S.P. Brown, *Probing hydrogen bonding in cocrystals and amorphous dispersions using <sup>14</sup>N-<sup>1</sup>H HMQC solid-state NMR*, *Mol. Pharmaceutics* 10 (2013) 999-1007.
71. P. Cerreia Vioglio, L. Catalano, V. Vasylyeva, C. Nervi, M.R. Chierotti, G. Resnati, R. Gobetto, P. Metrangolo, *Natural Abundance <sup>15</sup>N and <sup>13</sup>C Solid-State NMR Chemical Shifts: High Sensitivity Probes of the Halogen Bond Geometry*, *Chem. Eur. J.* 22 (2016) 16819-16828.
72. J. Viger-Gravel, S. Leclerc, I. Korobkov, D.L. Bryce, *Correlation between <sup>13</sup>C chemical shifts and the halogen bonding environment in a series of solid para-diiodotetrafluorobenzene complexes*, *CrystEngComm* 15 (2013) 3168-3177.
73. P.M.J. Szell, S.A. Gabriel, R.D. Gill, S.Y. Wan, B. Gabidullin, D.L. Bryce, *<sup>13</sup>C and <sup>19</sup>F solid-state NMR and X-ray crystallographic study of halogen-bonded frameworks featuring nitrogen-containing heterocycles*, *Acta Cryst.* C73 (2017) 157-167.
74. X. Yuan, T.X. Xiang, B.D. Anderson, E.J. Munson, *Hydrogen Bonding Interactions in Amorphous Indomethacin and Its Amorphous Solid Dispersions with Poly(vinylpyrrolidone) and Poly(vinylpyrrolidone-co-vinyl acetate) Studied Using <sup>13</sup>C Solid-State NMR*, *Mol. Pharmaceutics* 12 (2015) 4518-4528.
75. L. Rajput, M. Banik, J.R. Yarava, S. Joseph, M.K. Pandey, Y. Nishiyama, G.R. Desiraju, *Exploring the salt-cocrystal continuum with solid-state NMR using natural-abundance samples: implications for crystal engineering*, *IUCrJ* 4 (2017) 466-475.
76. H.E. Kerr, L.K. Softley, K. Suresh, P. Hodgkinson, I.R. Evans, *Structure and physicochemical characterization of a naproxen-picolinamide cocrystal*, *Acta Cryst.* C73 (2017) 168-175.
77. C.M. Widdifield, H. Robson, P. Hodgkinson, *Furosemide's one little hydrogen atom: NMR crystallography structure verification of powdered molecular organics*, *Chem. Commun.* 52 (2016) 6685-6688.
78. G.A. Facey, T.J. Connolly, C. Bensimon, T. Durst, *A solid state NMR and X-ray crystallographic investigation of dynamic disorder in solid tetrahydronaphthalene derivatives*, *Can. J. Chem.* 74 (1996) 1844-1851.
79. T. Venancio, L.M. Oliveira, T. Pawlak, J. Ellena, N. Boechat, S.P. Brown, *The use of variable temperature <sup>13</sup>C solid-state MAS NMR and GIPAW DFT calculations to explore the dynamics of diethylcarbamazine citrate*, *Magn. Reson. Chem.* 57 (2019) 200-210.
80. R.F. Moran, D.M. Dawson, S.E. Ashbrook, *Exploiting NMR spectroscopy for the study of disorder in solids*, *Int. Rev. Phys. Chem.* 36 (2017) 39-115.
81. H.E. Kerr, L.K. Softley, K. Suresh, A. Nangia, P. Hodgkinson, I.R. Evans, *A furosemide–isonicotinamide cocrystal: an investigation of properties and extensive structural disorder*, *CrystEngComm* 17 (2015) 6707-6715.
82. P. Florian, D. Massiot, *Beyond periodicity: probing disorder in crystalline materials by solid-state nuclear magnetic resonance spectroscopy*, *CrystEngComm* 15 (2013) 8623-8626.
83. E. Pindelska, A. Sokal, W. Kolodziejewski, *Pharmaceutical cocrystals, salts and polymorphs: Advanced characterization techniques*, *Adv. Drug Deliv. Rev.* 117 (2017) 111-146.
84. *BIOVIA Materials Studio 2017*. 2017, Dassault Systèmes.
85. *Topspin*. 2016, Bruker Biospin.
86. A.S. Tatton, L. Hughes, H. Blade, S.O.N. Lill, A. Bartók-Pártay, S.P. Brown, P. Hodgkinson. *Isolated molecule calculation protocol with Materials Studio*. 2019; Available from:

- <https://www.ccpnc.ac.uk/output/isolated-molecule-calculation-protocol-with-materials-studio>.
87. S. Sturniolo, T.F.G. Green, R.M. Hanson, M. Zilka, K. Refson, P. Hodgkinson, S.P. Brown, J.R. Yates, *Visualization and processing of computed solid-state NMR parameters: MagresView and MagresPython*, *Solid State Nucl. Magn. Reson.* 78 (2016) 64-70.
  88. CCP-NC database. 2021; Available from: <https://www.ccpnc.ac.uk/output/ccp-nc-database-of-nmr-data>.
  89. M.J. Duer, *Introduction to Solid-State NMR Spectroscopy*. 2005: Wiley-Blackwell.
  90. D.C. Apperley, R.K. Harris, P. Hodgkinson, *Solid-State NMR: Basic Principles and Practice*. 2012: Momentum Press.
  91. *Modern Methods in Solid-state NMR*. New Developments in NMR, ed. P. Hodgkinson. 2018: Royal Society of Chemistry.
  92. B. Reif, S.E. Ashbrook, L. Emsley, M. Hong, *Solid-state NMR spectroscopy*, *Nat. Rev. Dis. Primers* 1 (2021). DOI: 10.1038/s43586-020-00002-1.
  93. P. Hohenberg, W. Kohn, *Inhomogeneous Electron Gas*, *Phys. Rev.* 136 (1964) B864.
  94. W. Kohn, L.J. Sham, *Self-consistent equations including exchange and correlation effects*, *Phys. Rev.* 140 (1965) A1133-A1138.
  95. M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, J.D. Joannopoulos, *Iterative minimization techniques for ab initio total-energy calculations - molecular-dynamics and conjugate gradients*, *Rev. Mod. Phys.* 64 (1992) 1045-1097.
  96. Dassault Systèmes BIOVIA, *Materials Studio*, version 17, San Diego: Dassault Systèmes, 2017.
  97. J.P. Perdew, K. Burke, M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, *Phys. Rev. Lett.* 77 (1996) 3865-3868.
  98. A. Tkatchenko, M. Scheffler, *Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data*, *Phys. Rev. Lett.* 102 (2009) 073005.
  99. G.N. Reddy, D.S. Cook, D. Iuga, R.I. Walton, A. Marsh, S.P. Brown, *An NMR crystallography study of the hemihydrate of 2', 3'-O-isopropylidinediadenosine*, *Solid State Nucl. Magn. Reson.* 65 (2015) 41-48.
  100. J.D. Hartman, R.A. Kudla, G.M. Day, L.J. Mueller, G.J. Beran, *Benchmark fragment-based <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N and <sup>17</sup>O chemical shift predictions in molecular crystals*, *Phys. Chem. Chem. Phys.* 18 (2016) 21686-21709.
  101. S.T. Holmes, C.S. Vojvodin, R.W. Schurko, *Dispersion-Corrected DFT Methods for Applications in Nuclear Magnetic Resonance Crystallography*, *J. Phys. Chem. A* 124 (2020) 10312-10323.
  102. S.T. Holmes, O.G. Engl, M.N. Srnc, J.D. Madura, R. Quinones, J.K. Harper, R.W. Schurko, R.J. Luliucci, *Chemical Shift Tensors of Cimetidine Form A Modeled with Density Functional Theory Calculations: Implications for NMR Crystallography*, *J. Phys. Chem. A* 124 (2020) 3109-3119.
  103. B.G. Frommer, M. Côté, S.G. Louie, M.L. Cohen, *Relaxation of Crystals with the Quasi-Newton Method*, *J. Comput. Phys.* 131 (1997) 233-240.
  104. D.V. Dudenko, J.R. Yates, K.D.M. Harris, S.P. Brown, *An NMR crystallography DFT-D approach to analyse the role of intermolecular hydrogen bonding and  $\pi$ - $\pi$  interactions in driving cocrystallisation of indomethacin and nicotinamide*, *CrystEngComm* 15 (2013) 8797-8807.
  105. S. Adiga, D. Aebi, D.L. Bryce, *EFGShield — A program for parsing and summarizing the results of electric field gradient and nuclear magnetic shielding tensor calculations*, *Can. J. Chem.* 85 (2007) 496-505.
  106. D. Presti, A. Pedone, D. Licari, V. Barone, *A Modular Implementation for the Simulation of 1D and 2D Solid-State NMR Spectra of Quadrupolar Nuclei in the Virtual Multifrequency Spectrometer-Draw Graphical Interface*, *J. Chem. Theory Comput.* 13 (2017) 2215-2229.
  107. D. Licari, A. Baiardi, M. Biczysko, F. Egidi, C. Latouche, V. Barone, *Implementation of a graphical user interface for the virtual multifrequency spectrometer: The VMS-Draw tool*, *J. Comput. Chem.* 36 (2015) 321-34.

108. M. Marques, W. Braga, J.S. Junior, L. Prado, H. Rocha, F. Ferreira, G. Ferreira, J.A.L.C. Resende, *CCDC 1847168: Experimental Crystal Structure Determination*, CSD Communication (2019).
109. P.M.J. Szell, J.R. Lewandowski, H. Blade, L.P. Hughes, S.O. Nilsson Lill, S.P. Brown, *Taming the dynamics in a pharmaceutical by cocrystallization: investigating the impact of the cofomer by solid-state NMR*, *CrystEngComm* (2021). DOI: 10.1039/d1ce01084k.
110. A.S. Tatton, H. Blade, S.P. Brown, P. Hodgkinson, L.P. Hughes, S.O.N. Lill, J.R. Yates, *Improving Confidence in Crystal Structure Solutions Using NMR Crystallography: The Case of  $\beta$ -Piroxicam*, *Cryst. Growth Des.* 18 (2018) 3339-3351.
111. J. van de Streek, M.A. Neumann, *Validation of experimental molecular crystal structures with dispersion-corrected density functional theory calculations*, *Acta Crystallogr. Sect. B: Struct. Sci.* 66 (2010) 544-558.
112. J. van de Streek, M.A. Neumann, *Validation of molecular crystal structures from powder diffraction data with dispersion-corrected density functional theory (DFT-D)*, *Acta Crystallogr. Sect. B: Struct. Sci.* 70 (2014) 1020-1032.
113. A.E. Watts, K. Maruyoshi, C.E. Hughes, S.P. Brown, K.D.M. Harris, *Combining the Advantages of Powder X-ray Diffraction and NMR Crystallography in Structure Determination of the Pharmaceutical Material Cimetidine Hydrochloride*, *Cryst. Growth Des.* 16 (2016) 1798-1804.
114. C.M. Widdifield, J.D. Farrell, J.C. Cole, J.A.K. Howard, P. Hodgkinson, *Resolving alternative organic crystal structures using density functional theory and NMR chemical shifts*, *Chem. Sci.* 11 (2020) 2987-2992.
115. C.R. Groom, I.J. Bruno, M.P. Lightfoot, S.C. Ward, *The Cambridge Structural Database*, *Acta Crystallogr. Sect. B: Struct. Sci.* 72 (2016) 171-179.
116. E. Salager, G.M. Day, R.S. Stein, C.J. Pickard, B. Elena, L. Emsley, *Powder crystallography by combined crystal structure prediction and high-resolution  $^1\text{H}$  solid-state NMR spectroscopy*, *J. Am. Chem. Soc.* 132 (2010) 2564-2566.
117. M. Zilka, D.V. Dudenko, C.E. Hughes, P.A. Williams, S. Sturniolo, W.T. Franks, C.J. Pickard, J.R. Yates, K.D.M. Harris, S.P. Brown, *Ab initio random structure searching of organic molecular solids: assessment and validation against experimental data*, *Phys. Chem. Chem. Phys.* 19 (2017) 25949-25960.
118. R.K. Harris, P.Y. Ghi, R.B. Hammond, C.Y. Ma, K.J. Roberts, *Refinement of hydrogen atomic position in a hydrogen bond using a combination of solid-state NMR and computation*, *Chem. Commun.* (2003) 2834-2835.
119. R.K. Harris, P.Y. Ghi, R.B. Hammond, C.Y. Ma, K.J. Roberts, J.R. Yates, C.J. Pickard, *Solid-state NMR and computational studies of 4-methyl-2-nitroacetanilide*, *Magn. Reson. Chem.* 44 (2006) 325-333.
120. M.G. Siskos, M.I. Choudhary, I.P. Gerathanassis, *Refinement of labile hydrogen positions based on DFT calculations of  $^1\text{H}$  NMR chemical shifts: comparison with X-ray and neutron diffraction methods*, *Org. Biomol. Chem.* 15 (2017) 4655-4666.
121. L. Wang, F.J. Uribe-Romo, L.J. Mueller, J.K. Harper, *Predicting anisotropic thermal displacements for hydrogens from solid-state NMR: a study on hydrogen bonding in polymorphs of palmitic acid*, *Phys. Chem. Chem. Phys.* 20 (2018) 8475-8487.
122. A.L. Reviglio, F.A. Martínez, M.D.A. Montero, Y. Garro-Linck, G.A. Aucar, N.R. Sperandeo, G.A. Monti, *Accurate location of hydrogen atoms in hydrogen bonds of tizoxanide from the combination of experimental and theoretical models*, *RSC Adv.* 11 (2021) 7644-7652.
123. L.J. Cui, X.Y. Chen, *Isonicotinonitrile-benzoic acid (1/1)*, *Acta Crystallogr Sect E Struct Rep Online* 66 (2010) o722.
124. A. Gaywood, P.A. Wood, H. McNab, *CCDC 1412517: Experimental Crystal Structure Determination*, CSD Communication (2015).
125. J.R. Yates, T.N. Pham, C.J. Pickard, F. Mauri, A.M. Amado, A.M. Gil, S.P. Brown, *An investigation of weak CH...O hydrogen bonds in maltose anomers by a combination of*

- calculation and experimental solid-state NMR spectroscopy*, J. Am. Chem. Soc. 127 (2005) 10216-10220.
126. C. Gervais, M. Profeta, V. Lafond, C. Bonhomme, T. Azais, H. Mutin, C.J. Pickard, F. Mauri, F. Babonneau, *Combined ab initio computational and experimental multinuclear solid-state magnetic resonance study of phenylphosphonic acid*, Magn. Reson. Chem. 42 (2004) 445-452.
127. J. Schmidt, A. Hoffmann, H.W. Spiess, D. Sebastiani, *Bulk chemical shifts in hydrogen-bonded systems from first-principles calculations and solid-state-NMR*, J. Phys. Chem. B 110 (2006) 23204-23210.
128. M. Zilka, S. Sturniolo, S.P. Brown, J.R. Yates, *Visualising crystal packing interactions in solid-state NMR: Concepts and applications*, J. Chem. Phys. 147 (2017) 144203.
129. A.C. Uldry, J.M. Griffin, J.R. Yates, M. Perez-Torralba, M.D. Maria, A.L. Webber, M.L. Beaumont, A. Samoson, R.M. Claramunt, C.J. Pickard, et al., *Quantifying weak hydrogen bonding in uracil and 4-cyano-4'-ethynylbiphenyl: a combined computational and experimental investigation of NMR chemical shifts in the solid state*, J. Am. Chem. Soc. 130 (2008) 945-954.
130. B.L. Dale, N.R. Halcovitch, M.J.G. Peach, J.M. Griffin, *Investigation of structure and dynamics in a photochromic molecular crystal by NMR crystallography*, Magn. Reson. Chem. 57 (2019) 230-242.
131. A.C. Poppler, E.K. Corlett, H. Pearce, M.P. Seymour, M. Reid, M.G. Montgomery, S.P. Brown, *Single-crystal X-ray diffraction and NMR crystallography of a 1:1 cocrystal of dithianon and pyrimethanil*, Acta Cryst. C73 (2017) 149-156.
132. G.N.M. Reddy, A. Marsh, J.T. Davis, S. Masiero, S.P. Brown, *Interplay of Noncovalent Interactions in Ribbon-like Guanosine Self-Assembly: An NMR Crystallography Study*, Cryst. Growth Des. 15 (2015) 5945-5954.
133. V.K. Srirambhatla, A. Kraft, S. Watt, A.V. Powell, *A robust two-dimensional hydrogen-bonded network for the predictable assembly of ternary co-crystals of furosemide*, CrystEngComm 16 (2014) 9979-9982.
134. See: <https://creativecommons.org/licenses/by/4.0/>.