

# Lawrence Berkeley National Laboratory

LBL Publications

## Title

Bridging experiment and theory: a template for unifying NMR data and electronic structure calculations

## Permalink

<https://escholarship.org/uc/item/3966r42k>

## Journal

Journal of Cheminformatics, 8(1)

## ISSN

1758-2946

## Authors

Brown, David ML

Cho, Herman

de Jong, Wibe A

## Publication Date

2016-12-01

## DOI

10.1186/s13321-016-0120-z

Peer reviewed

METHODOLOGY

Open Access



# Bridging experiment and theory: a template for unifying NMR data and electronic structure calculations

David M. L. Brown<sup>1†</sup>, Herman Cho<sup>2\*†</sup> and Wibe A. de Jong<sup>3†</sup>

## Abstract

**Background:** The testing of theoretical models with experimental data is an integral part of the scientific method, and a logical place to search for new ways of stimulating scientific productivity. Often experiment/theory comparisons may be viewed as a workflow comprised of well-defined, rote operations distributed over several distinct computers, as exemplified by the way in which predictions from electronic structure theories are evaluated with results from spectroscopic experiments. For workflows such as this, which may be laborious and time consuming to perform manually, software that could orchestrate the operations and transfer results between computers in a seamless and automated fashion would offer major efficiency gains. Such tools also promise to alter how researchers interact with data outside their field of specialization by, e.g., making raw experimental results more accessible to theorists, and the outputs of theoretical calculations more readily comprehended by experimentalists.

**Results:** An implementation of an automated workflow has been developed for the integrated analysis of data from nuclear magnetic resonance (NMR) experiments and electronic structure calculations. Kepler (Altintas et al. 2004) open source software was used to coordinate the processing and transfer of data at each step of the workflow. This workflow incorporated several open source software components, including electronic structure code to compute NMR parameters, a program to simulate NMR signals, NMR data processing programs, and others. The Kepler software was found to be sufficiently flexible to address several minor implementation challenges without recourse to other software solutions. The automated workflow was demonstrated with data from a <sup>17</sup>O NMR study of uranyl salts described previously (Cho et al. in *J Chem Phys* 132:084501, 2010).

**Conclusions:** The functional implementation of an automated process linking NMR data with electronic structure predictions demonstrates that modern software tools such as Kepler can be used to construct programs that comprehensively manage complex, multi-step scientific workflows spanning several different computers. Automation of the workflow can greatly accelerate the pace of discovery, and allows researchers to focus on the fundamental scientific questions rather than mastery of specialized software and data processing techniques. Future developments that would expand the scope and power of this approach include tools to standardize data and associated metadata formats, and the creation of interactive user interfaces to allow real-time exploration of the effects of program inputs on calculated outputs.

**Keywords:** Scientific workflow, NMR spectroscopy, Electronic structure theory

\*Correspondence: hm.cho@pnl.gov

<sup>†</sup>David M. L. Brown, Herman Cho, and Wibe A. de Jong contributed equally to this work

<sup>2</sup>Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA

Full list of author information is available at the end of the article

## Background

In the physical sciences, a complex series of steps is often required to relate a theoretical hypothesis to an experimental observable, and *vice versa*. The study of electronic structure by nuclear magnetic resonance (NMR) spectroscopy illustrates the difficulties that can arise with this process. The object of this particular workflow is to transform results from electronic structure simulations into predicted NMR spectra or, in reverse, to extract electronic structure parameters from observed energies and line shapes. This practice can be found in some of the earliest accounts of NMR spectroscopy [1, 2], and it continues to be a valuable and popular approach for elucidating the electronic structure of molecules and crystals.

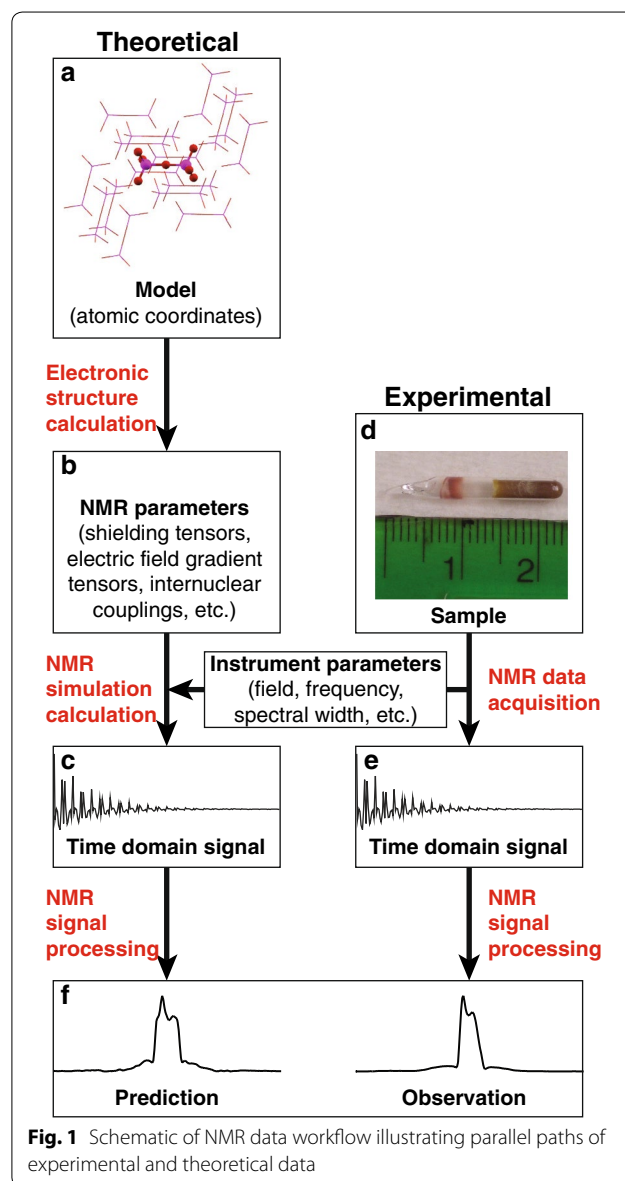
A schematic of the forward transformation is shown in Fig. 1. As portrayed in this figure, the workflow encompasses an array of independent computer programs and data inputs, each requiring specialized knowledge for their use. Manual step-wise execution of this workflow is a cumbersome process ill-suited for efficient, interactive fitting of theoretical models and experimental data. Automation of the intermediate steps would greatly expedite the workflow, but in practice requires the merging of software and data from a multitude of instrument makers and electronic structure codes. A further complication is the large variety of NMR experiments and observables an automated workflow might need to accommodate, which necessitates the compilation of a library of experiment-specific simulation programs.

In this paper, we demonstrate how the NMR workflow can be consolidated and simplified with the use of software tools that execute intermediate operations automatically and invisibly. This implementation is part of an effort to enhance the interactivity of experimentalists and theorists that we refer to as the EMSL Experiment/Theory Unification Project (EETUP) [3]. A flexible and modular software architecture has been developed that can accommodate a diverse set of open source software packages, and allows the range of functionality to be expanded with additional modules as new requirements arise. Initial development efforts have been focused on the analysis of NMR spectra of quadrupolar nuclides, which provide measurements of both chemical shifts and electric field gradients, but more universal applications are possible through the addition of other spectral simulation modules.

## Methods

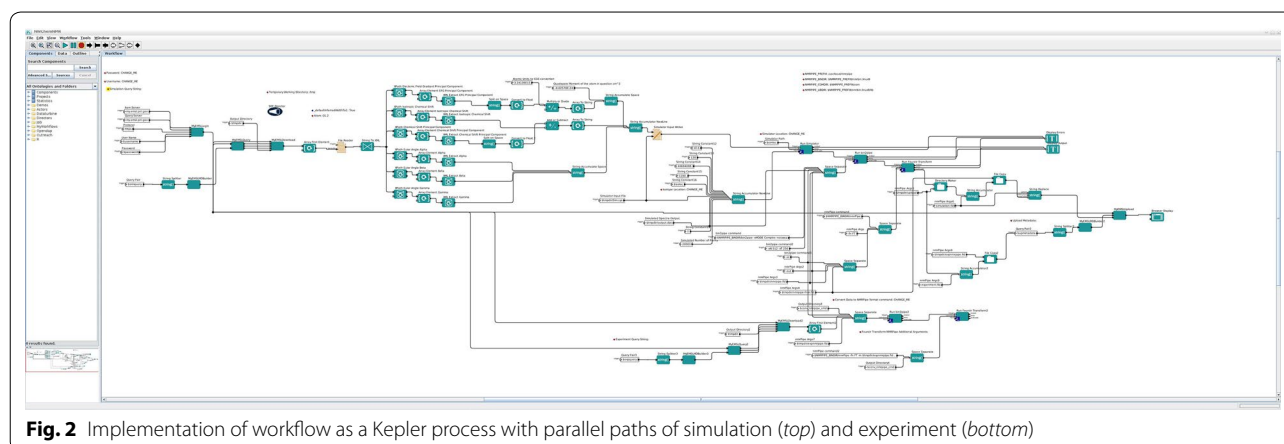
### Workflow control

The Kepler Project [4] offers software tools designed to orchestrate complex scientific workflows [5]. In particular, the Kepler interface allows users to create workflow



control programs without explicitly writing source code. In addition, Kepler enforces good coding practices in workflow design, including modularity and extensibility of software. The master process we have constructed to manage the workflow in Fig. 1 has been assembled out of tools created by Kepler Project developers. A step-by-step representation of the master Kepler process appears in Fig. 2

The workflow typically spans several different platforms, each dedicated to a specific task: experimental data are acquired with one computer, electronic structure parameters are calculated on another computer, NMR spectra are simulated on a third computer, and so



**Fig. 2** Implementation of workflow as a Kepler process with parallel paths of simulation (*top*) and experiment (*bottom*)

forth. With no single computer controlling the overall workflow, it proves critical to store data in a centralized location. The Active Data Library at our site, MyEMSL, serves as the central access point for EETUP processes. In addition, MyEMSL provides application programming interfaces (APIs) for authentication, querying, and transfers of data. The APIs were abstracted to Kepler Actors through the system SWADL [6–8].

The master Kepler process was programmed in accordance with standard practices to ensure portability to other platforms and adaptability to future needs. Maximal use of actors native to Kepler was made to support essential functions, and a gap analysis was performed to ensure proper exception handling, either with a native actor [9] or an *ExternalExecutionEnvironmentActor* (triple-E) actor that executes custom binary code.

The entry point for the theoretical side of the workflow in Fig. 1a is the calculation of NMR parameters. Electronic structure simulation software relies on sophisticated decisions about basis sets, functionals, molecular structures, atom selections, etc., to be utilized effectively. Due to the specialized and fluid nature of electronic structure codes, the outputs of these programs frequently require checking both for physical reasonableness and for compatibility with subsequent programs in the workflow chain. At present, the complexity of these operations requires direct human interaction with the software and precludes automation within Kepler.

Subsequent steps of the workflow are more readily compatible with automation. The automated part of the workflow in the current implementation begins with the NMR spectral simulation. The simulation code exists as a distinct standalone executable program. Kepler file writer actors prepare the inputs for the code, the triple-E actor executes the simulations, and file reader actors direct the simulation output to the next step of the workflow. The experimental (Fig. 1e) and simulated (Fig. 1c) data are

passed through the same NMR signal processing application, and processed in an equivalent manner.

Inputs for the spectral simulation step are entered into the Kepler process as strings, since they are ultimately passed as command line or text input files to the programs performing the calculations. The Kepler process can verify input by converting the values to their appropriate types then back to strings.

#### Workflow software and data formats

An array of software choices is available for each step of the workflow. Table 1 lists the applications that were chosen for our current implementation. The applications in this table are open source, in widespread use, and readily extensible. Alternative choices at each step of the workflow can readily be incorporated as user-selectable options within the modular Kepler framework. The software at each step and their input and output data formats are described below.

#### Electronic structure calculation

The electronic structure code used in this implementation, NWChem [10], accepts text inputs and generates

**Table 1** Software used in example NMR workflow (refer to Fig. 1)

Task	Program name	Source
Electronic structure calculation	NWChem	Valiev et al. [10]
NMR spectral simulation	Gamma	Smith et al. [11]
Simulated NMR signal processing	NMRPipe	Delaglio et al. [12]
NMR experimental data acquisition	NTNMR	Tecmag, Inc.
	VnmrJ	Agilent
	TopSpin	Bruker
Experimental NMR signal processing	NMRPipe	Delaglio et al. [12]

Chemical Markup Language (CML) output [13, 14], which is stored on MyEMSL (Fig. 1b). Since CML is a subset of XML, standard Kepler actors for parsing XML are able to extract the data required by the simulation. The use of text files for input and output facilitates the human interaction needed to execute electronic structure codes and interpret results.

Analyses of electronic structure in general require expert decisions by the user on the portion of a molecule or lattice that is to be included in the computation of electronic structure and NMR coupling parameters. This task is performed by manual entry of the coordinates of the selected atoms into the relevant programs, although a graphical user interface could readily be conceived to perform this operation more conveniently and reliably.

#### NMR instrument parameters

NMR instrument data are typically stored in files with proprietary binary formats unique to the instrument's manufacturer. The C++ programs created by us parse current generation data files from NMR instruments manufactured by Agilent, Bruker, and Tecmag. Automated execution of these standalone programs was handled by a triple-E actor in Kepler.

#### Spectral simulation inputs and outputs

The predicted NMR signal is computed from the inputs in Table 2 as a time domain interferogram using the time-dependent density matrix formalism [2]. Simulated spectra in the current implementation are produced by custom C++ programs linked against the GAMMA version 4.1.0 NMR simulation environment [11]. Required input data for the simulation are listed in Table 2. Electronic structure and NMR instrument data are obtained

as outlined above. The nuclear parameters such as gyromagnetic ratios and quadrupolar couplings represent a relatively small amount of static data, and may be compiled from reference databases and saved in a text file for ease of reading and updating.

In addition to the data displayed in Table 2, geometric parameters specifying the orientation of the tensor principal axes with respect to the applied magnetic field direction must be supplied. These data can be in the form of a longitudinal and azimuthal pair of angles representing a single configuration of the tensors, or more commonly, an array of angle pairs representing multiple orientations of the tensors to model a disordered ensemble of nuclear spin systems. To accommodate different models, from a single orientation to an ensemble, the simulation program reads the geometric data from a file and computes and adds spectra for all of the orientations contained in the file.

For correct alignment, the experimental and simulated spectra must be centered at the same chemical shift value, and have equal digital resolutions. The automated process we use to perform the alignment of spectra is explained in Appendix 1.

At present, the outputs of the simulation calculation are stored in a binary packed format directly readable by the processing software selected for our implementation, viz., NMRPipe.

#### Signal processing and visualization

Both the experimental (Fig. 1e) and simulated (Fig. 1c) results are in the form of time domain data, and require processing to obtain frequency domain spectra (Fig. 1f). We have chosen the NMRPipe [12] software package for our first effort to integrate automated data processing in the workflow. NMRPipe is an attractive choice for several reasons: it is open source software in wide use in the NMR community, and provides a comprehensive set of NMR data processing tools. Data analysis in NMRPipe is separated from data display, which greatly simplifies the integration with other processes unconnected with data analysis, such as data uploading and orchestration.

NMRPipe recognizes the data formats of all of the major NMR instrument manufacturers, eliminating the need to create programs to translate data to a readable form. Data are passed between individual NMRPipe processes via pipes (see Fig. 3), reading and writing from standard input (*stdin*) and standard output (*stdout*) streams, respectively. Input analysis parameters are entered from command line arguments. To automate this process the *stdout* output stream from one triple-E actor was passed to the *stdin* input stream of the subsequent triple-E actor. Triple-E actors are unable to directly pass data in the preferred format of NMRPipe processes

**Table 2 Data entered into workflow simulation program**

Source	Parameter
Electronic structure calculation	Shielding tensor principal values ( $\sigma_{ij}$ )
	Electric field gradient parameters ( $V_{zz}, \eta_Q$ )
	Euler angles relating principal axis systems of $\sigma$ and $\mathbf{V}$ ( $\alpha, \beta, \gamma$ )
	Isotropic shielding value of chemical shift reference ( $\bar{\sigma}_{ref}$ )
NMR instrument data file	Spectrometer carrier frequency ( $\nu$ )
	Frequency at 0 ppm ( $\nu_0$ )
	Spectral digital resolution
Nuclear parameter database	Gyromagnetic ratio ( $\gamma$ )
	Nuclear spin quantum number ( $I$ )
	Quadrupole moment ( $Q$ )
Molecular structure	Atomic coordinates ( $\vec{r}_i$ )
	Internuclear distances and vectors ( $\vec{r}_{jk}$ )

```

#!/bin/csh

bin2pipe -in ./UranylAmmCarbStructC.dat -xN 512 -xT 256 -xOBS
40.739 -xSW 1000000 -xMODE Complex -noswap \
| nmrPipe -fn EM -lb 10000 \
| nmrPipe -fn ZF -auto \
| nmrPipe -fn FT \
| nmrPipe -fn PS -p0 0 -p1 0.0 -di \
| nmrPipe -fn POLY -auto \
| nmrPipe -fn EXT -x1 10000ppm -xn -10000ppm -sw \
-out simulation.fid -verb 2 -ov

#!/bin/csh

rm -f experimental.fid
bin2pipe -bo 1056 -in ./032606a.tnt -xN 512 -xT 256 -xOBS 40.739 -
xSW 1000000 -xMODE Complex -noswap \
| nmrPipe -fn SP -off 0.0 -end 1.0 -pow 2.0 \
| nmrPipe -fn FT \
| nmrPipe -fn PS -p0 -58.0 -p1 46080.0 -di \
| nmrPipe -fn EXT -x1 10000ppm -xn -10000ppm -sw \
-out experimental.fid

```

**Fig. 3** NMRPipe shell scripts used to process simulated (*top*) and experimental (*bottom*)  $^{17}\text{O}$  NMR data of  $(\text{NH}_4)_4\text{UO}_2(\text{CO}_3)_3$ , with outputs as shown in Fig. 5. These scripts were integrated into the Kepler workflow

(compressed binary) necessitating the creation of temporary files as the intermediary of data transfers between processes.

The NMRPipe tool, NMRView, is used for final data visualization.

### Case study

A recently published solid-state  $^{17}\text{O}$  NMR study of  $^{17}\text{O}$ -enriched uranyl salts serves to illustrate the performance of the automated workflow [15]. In this case, experimental results were acquired on a Tecmag, Inc., NMR spectrometer controlled by a computer running a Windows XP operating system. Files stored on this computer were transferred to the centralized data repository, MyEMSL, along with the outputs of the electronic structure calculation performed on the EMSL high performance computer [16].

Upon completion of the data uploads to MyEMSL the automated part of the workflow was initiated on a desktop computer executing the master Kepler process via the scripts shown in Fig. 4. The spectral simulation and processing of the experimental and simulated time domain data were performed on this computer with no further human intervention, and the results directed to MyEMSL (steps  $B \rightarrow C \rightarrow F$  and  $E \rightarrow F$  in Fig. 1). The final result displayed by the desktop machine appears in Fig. 5, which shows a screen capture of the NMRView window with the predicted

(top) and actual (bottom)  $^{17}\text{O}$  NMR spectra. These spectra may be compared to Figure 5 of reference [15].

The pace of this workflow is slowed by user intervention; the task for the computer at each step may be completed within seconds, but manual data entry and program execution might require several hours of concentrated human effort. By automatically streaming data from computer to computer the Kepler process eliminates the tedious manual steps and can transform the workflow into an instantly interactive operation.

### Conclusion

The implementation described here can serve as a template for the automation of other workflows that blend experimental observables and computational theory. Tools from the Kepler Project provide the capability that allows multiple platforms running sophisticated standalone software to be merged and executed with minimal intervention or expert knowledge on the part of the user. Theory results are made more accessible to experimentalists, and experimental data are more readily interpreted by theorists. All software and documentation developed to date are publicly accessible [3]. Future releases and updates will be made available at this same site. The custom Kepler actors created for this project are also provided at these sites [7], but have not yet been accepted as part of the official Kepler release.

```

#!/bin/bash

USERNAME=$1
PASSWORD=""
read -s -p "PASSWORD:" PASSWORD

kepler -runkar \
  -myemsl_username "$USERNAME" \
  -myemsl_password "$PASSWORD" \
  -simquery 'proposal=45796#NWChem.CML=test_apr_17' \
  -expquery 'proposal=45796#Tag=NMR_EXP_jun_19' \
  -upmetadata 'proposal=45796#Tag=nmrpipe_test_jun_19' \
  -simloc /home/dmlb2000/nmrtools/trunk/src/sim/
    Quad_NMR_Simulator \
  -isoloc /home/dmlb2000/nmrtools/trunk/src/sim/oxygen17.inp.
    txt \
  -conv_nmrpipe_cmd '/usr/local/nmrpipe/nmrbin.linux9/
    bin2pipe -bo 1056 -xN 256 -xT 128 -xMODE Complex -
    noswap' \
  $PWD/NWChemNMR.kar

#!/bin/bash

USERNAME=$1
PASSWORD=""
read -s -p "PASSWORD:" PASSWORD

MY_TMPDIR=$(mktemp -d)

/home/dmlb2000/kepler-nobuild/kepler-2.4/kepler.sh -runkar \
  -myemsl_username "$USERNAME" \
  -myemsl_password "$PASSWORD" \
  -tmpdir "$MY_TMPDIR" \
  -NMRPIPE_PREFIX '/home/dmlb2000/nmr' \
  -query_server 'a3.my.emsl.pnl.gov' \
  -data_server 'a4.my.emsl.pnl.gov' \
  -simquery 'proposal=45796#NWChem.CML=test_apr_17' \
  -expquery 'proposal=45796#Tag=NMR_EXP_jun_19' \
  -upmetadata 'proposal=45796#Tag=nmrpipe_test_jun_19' \
  -simloc /home/dmlb2000/svn-repos/nmrtools/trunk/src/sim/
    Quad_NMR_Simulator \
  -isoloc /home/dmlb2000/svn-repos/nmrtools/trunk/src/sim/
    oxygen17.inp.txt \
  -conv_nmrpipe_cmd '/home/dmlb2000/nmr/nmrbin.linux9/
    bin2pipe -bo 1056 -xN 256 -xT 128 -xMODE Complex -
    noswap' \
  $PWD/NWChemNMR.kar

echo $MY_TMPDIR
ls $MY_TMPDIR

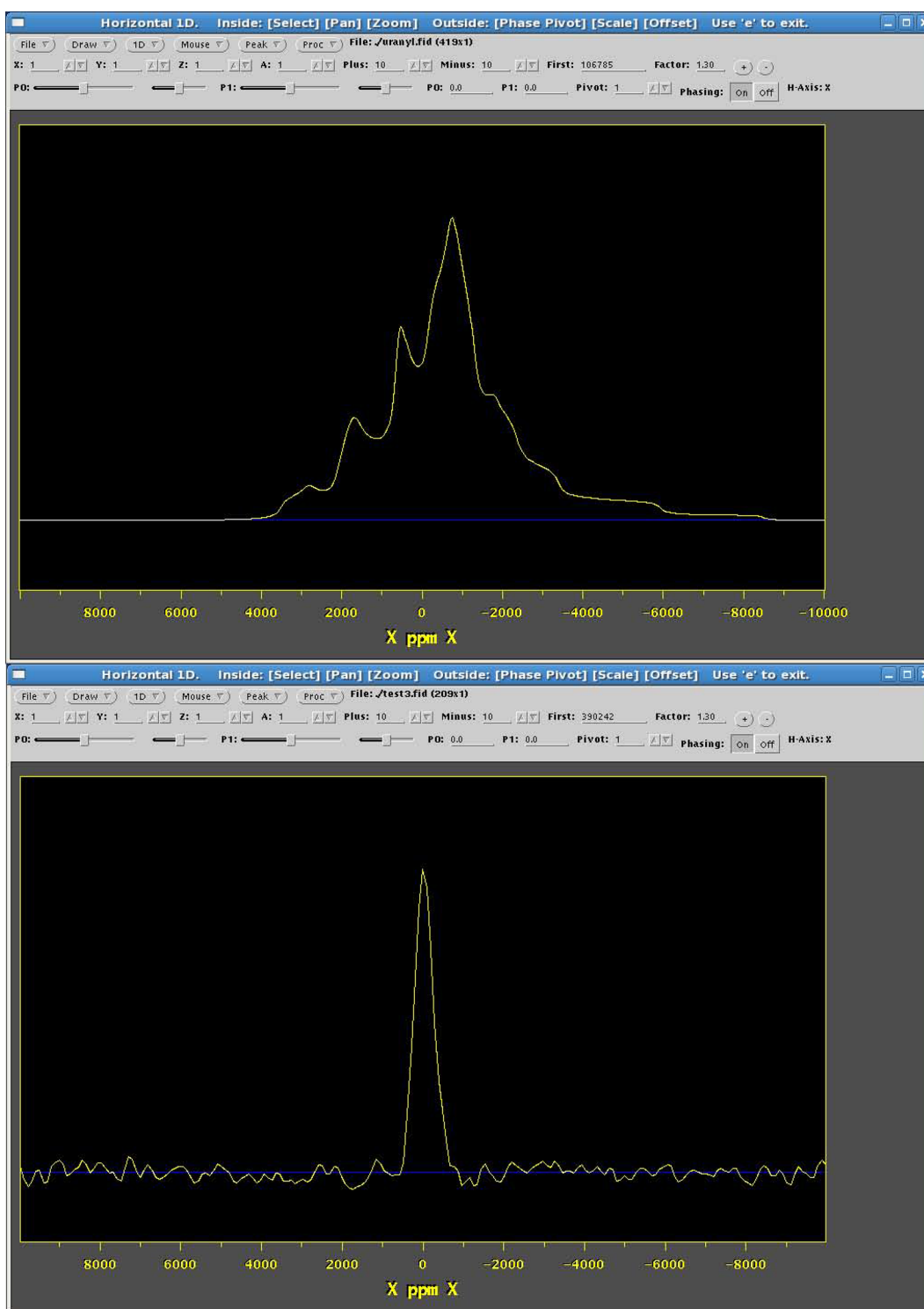
```

**Fig. 4** Two scripts for launching Kepler from the operating system command line for convenient analysis when new simulation or experimental results became available

The value of workflow tools will depend to a large extent on their scope and versatility, and in particular their ability to assimilate and process inputs from a wide

range of different sources at each step of the workflow. In our current implementation, we have created specialized software tools to read the data formats of the programs





**Fig. 5** NMRDraw output of simulated (*top*) and experimental (*bottom*)  $^{17}\text{O}$  NMR spectra of  $(\text{NH}_4)_4\text{UO}_2(\text{CO}_3)_3$



in Table 1, but the programming effort and complexity would rapidly increase as more choices were added to the selection in this table. It is clear that expandability of the workflow would be greatly facilitated if data files were standardized to make them universally readable. Standardization of data formats has not been widely implemented [17, 18], but even if adopted at a limited, local level a single unified data format can significantly simplify workflow development.

This implementation would be further improved by automating the creation of NWChem inputs from, e.g., molecular structure data, and starting the NWChem process. Software offered by Avogadro [19] may be superior to Kepler products in this regard and is under consideration as the path for future enhancement. While we foresee no fundamental obstacle to adding this functionality, the specialized knowledge required to select reasonable parameters and estimate computer resources make this a more difficult programming challenge than the ones considered thus far.

Although a central goal of EETUP is the seamless, automatic bridging of theoretical and experimental data, the ability to interrupt and manipulate inputs to the workflow at intermediate steps would add valuable new functionality. Real-time updating of a spectrum as a bond distance or shielding tensor is varied is one conceivable way where such a capability might enhance experiment/theory interactivity.

#### Authors' contributions

DMLB created, tested, and ran the Kepler workflow. HC created the spectral simulation and data translation software. All authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup> Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352, USA. <sup>2</sup> Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA. <sup>3</sup> Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

#### Acknowledgements

The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy. This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Basic Energy Sciences, Heavy Element Chemistry program. A portion of the research was performed using EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory.

#### Competing interests

The authors declare that they have no competing interests.

## Appendix 1: Alignment of experimental and simulated NMR spectra

The automated alignment of experimental and simulated NMR spectra requires a well-defined procedure for relating the field dependent frequency shift computed by electronic structure codes, viz., the nuclear shielding, with the chemical shift measured in NMR experiments. The

chemical shift is defined as the scaled difference from the position of a reference signal [20]

$$\delta_{\text{sample}} = (\nu_{\text{sample}} - \nu_{\text{ref}}) / \nu_{\text{ref}}, \quad (1)$$

thus to obtain a chemical shift  $\delta_{\text{sample}}$  from a computed shielding result, shielding tensors for the reference compound *and* the compound of interest must both be calculated. Reference compounds for stable NMR-active nuclides have been recommended by the IUPAC [20]. Outputs from two electronic structure calculations are therefore needed by the simulation code, although only the mean of the shielding tensor principal values  $\bar{\sigma}_{\text{ref}}$  is utilized from the calculation for the reference compound. These translations were implemented using standard Kepler math actors.

Correct alignment of the simulated spectrum with the experimental result is achieved by specifying simulation parameters that ensure the two spectra have equal digital resolution and the same chemical shift values at the spectrum center. The spectral width need not match, and indeed it can be desirable to have the simulation with a wider frequency range to display features that are not seen in experimental spectrum due to instrument limitations. Spectral alignment exemplifies an operation that involves the interplay of multiple different processes and computers, in this case the instrument computer, the simulation computer, and the machine that performs the electronic structure calculation. Kepler tools were used here for their ability to automate formulaic tasks across platforms, allowing global workflow parameters to be created, manipulated with mathematical actors, and made accessible throughout the workflow.

Received: 13 August 2015 Accepted: 27 January 2016

Published online: 09 February 2016

#### References

1. Abragam A (1961) The principles of nuclear magnetism. Clarendon Press, Oxford
2. Slichter CP (1990) Principles of magnetic resonance, 2nd edn. Springer, New York
3. Cho H, Brown DML Jr EMSL experiment/theory unification project. <https://github.com/dmlb2000/EETUP>
4. Altintas I, Berkley C, Jaeger E, Jones M, Ludascher B, Mock S (2004) Kepler: an extensible system for design and execution of scientific workflows. In: Proceedings of the 16th international conference on scientific and statistical database management, 2004, pp 423–424
5. Borreguero JM, Campbell SJ, Delaire OA, Doucet M, Goswami M, Hagen ME, Lynch VE, Proffen TE, Ren S, Savici AT et al (2014) Integrating advanced materials simulation techniques into an automated data analysis workflow at the spallation neutron source. In: TMS 2014 143rd annual meeting and exhibition, annual meeting supplemental proceedings. John Wiley Sons, p 297
6. Brown DML Jr (2015) A generic interface between scientific workflow tools and active data libraries. Master's thesis, Washington State University, Pullman

- Brown DML Jr SWADL Kepler components. <https://github.com/dmlb2000/kepler-swadl>
- Brown DML Jr Scientific workflow for active data libraries. <https://github.com/dmlb2000/swadl-library>
- Kepler example actor tutorial. <https://kepler-project.org/developers/teams/build/documentation/developing-a-hello-world-actor-using-the-kepler-build-system-and-eclipse>
- Valiev M, Bylaska EJ, Govind N, Kowalski K, Straatsma TP, Van Dam HJ, Wang D, Nieplocha J, Apra E, Windus TL et al (2010) NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput Phys Commun* 181(9):1477–1489
- Smith SA, Levante TO, Meier BH, Ernst RR (1994) Computer simulations in magnetic resonance. An object-oriented programming approach. *J Magn Reson A* 106(1):75–105
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293
- De Jong WA Modified NWChem with CML support. <https://github.com/dmlb2000/nwchem-cml>
- de Jong WA, Walker AM, Hanwell MD (2013) From data to analysis: linking NWChem and Avogadro with the syntax and semantics of Chemical Markup Language. *J Cheminform* 5:25
- Cho H, De Jong WA, Soderquist CZ (2010) Probing the oxygen environment in  $\text{UO}_2^{+2}$  by solid-state  $^{17}\text{O}$  nuclear magnetic resonance spectroscopy and relativistic density functional calculations. *J Chem Phys* 132:084501
- Environmental Molecular Sciences Laboratory: Molecular Science Computing. <https://www.emsl.pnl.gov/emslweb/capabilities/computing>
- McDonald RS, Wilks PA Jr (1988) JCAMP-DX a standard form for exchange of infrared spectra in computer readable form. *Appl Spectrosc* 42(1):151–162
- Davies AN, Lampen P (1993) JCAMP-DX for NMR. *Appl Spectrosc* 47(8):1093–1099
- Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminform* 4(1):17
- Harris RK, Becker ED, De Menezes SC, Goodfellow R, Granger P (2001) NMR nomenclature. Nuclear spin properties and conventions for chemical shifts (IUPAC recommendations, 2001). *Pure Appl Chem* 73(11):1795–1818

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---