

www.nmrdb.org: Resurrecting and Processing NMR Spectra On-line

Damiano Banfi[§] and Luc Patiny*

[§]SCS Poster Prize Winner

Abstract: NMR spectroscopy is certainly the analytical methodology that provides the most information about a molecule but the long-term vision of this information is often lacking. We present here two innovative tools accessible free from the web. NMR assigner allows a chemical structure to be assigned to the corresponding NMR spectrum by simply drawing lines between atoms and automatically characterized signals. NMR resurrector allows NMR spectra to be recreated from published in-line experimental parts enabling the recovery of this lost knowledge.

Keywords: ELN · ERM · LIMS · NMR spectrum assignment · On-line processing

NMR spectroscopy is probably the analytical methodology that provides the most information about a molecule. While acquiring a NMR spectrum is an easy or even automatic process, long-term vision on the information present in the spectrum may be difficult to foresee and we need to consider essential questions. Will I be able to read my spectra in ten years? How can spectra be compared? What about all my old assignments? How can molecules containing spectra NMR characteristics be found?

While most NMR spectra acquired up to now only exist as a string published in the literature and seem to be lost, new acquisitions must be saved and assigned in order to guarantee not only the endurance of the information but also the possibility to retrieve it and to use it for data mining. In 1988, a text format called 'JCAMP-DX' was created in order to save infrared spec-

tral data.^[1] Later this format was extended to store mass spectra,^[2] NMR spectra and even 2D NMR spectra.^[3,4] While this format is not optimized for its compactness, its published description and its validation by IUPAC is a guarantee of long-term availability. Moreover, this format is a standard exportation and importation format from many programs such as TopSpin.^[5]

On the other hand, while biologists are used to share their results for free on the internet this behavior for chemists is rather exceptional and most of the databases are expensive. A few databases containing NMR spectra are available for free on the internet like NmrShiftDB,^[6] the ChemExper Chemical Directory^[7] and the Spectral Database for Organic Compounds.^[8]

Not losing any information should be the leitmotif of today's chemist and to reach this goal we present here two new tools accessible for free on the internet: *NMR assigner*, to assign NMR spectra and *NMR resurrector*, to recreate a NMR spectrum from a NMR string assignment. Those tools are available from any web-browser with java virtual machine 1.4+ installed.^[9]

NMR Assigner: Assigning NMR Spectra

Assigning NMR spectra can be a time-consuming and apparently unrewarding process. As a result, there is a tendency to limit the analysis to a minimum and fulfill only short-term needs. Furthermore, data obtained from these not fully assigned spectra is very seldom stored in a conve-

niently structured database. This leads to a situation where a considerable amount of potentially valuable information is lost. For this reason it is very rare in both academic and corporate environments to be able to query a database for all singlets at 3 ppm corresponding to protons next to a carbonyl. Streamlining the process by removing as many of the repetitive and error-prone tasks as possible is a necessary step towards such a repository. Programs to process the NMR spectra and even store the results exist but most of them are either expensive (*i.e.* ACD/SpecManager,^[10] MestRe-C,^[11] iNMR^[12]) or have been designed for protein structure elucidation and are not really user friendly (*i.e.* NMRPipe^[13]).

While not intended as a replacement for the spectroscopist, this on-line tool helps by increasing efficiency in daily tasks.

The assignment of a NMR spectrum can be broken down to iterations of the following three steps (Fig. 1):

- Identification of the signal
- Integration, chemical shift, multiplicity and coupling constants determination
- Assignment of each signal to the corresponding atom in the molecule

In order to make this process straightforward, this tool includes a 'smart peak picking' function that allows each signal to be fully characterized with a single mouse movement. The structural assignment can then be performed simply by drawing a line between the 'smart peak label' and the corresponding atom.

Once the entire molecule has been processed, the assignment can be exported by a click of a button either in the 'ACS' format

*Correspondence: Dr. L. Patiny
Ecole Polytechnique Fédérale de Lausanne
EPFL - BCH 5121
CH-1015 Lausanne
Tel.: +41 21 693 94 69
E-mail: luc.patiny@epfl.ch

Bruker TopSpin
 * Spectrum acquisition
 * Fourier transform: **ft**
 * Phase correction: **apk**
 * Baseline correction: **abs**
 * Convert to jcamp-dx: **tojdx**
 Output data: SPEC REAL + IMAG
 Region file: Region file not used
 Compression mode: DIFF/DUP

```

##TITLE=B1284/010/nizatine RED5149
##JCAMPDX= 6.0          $$ Bruker NMR JCAMP-DX V2.0
##DATA TYPE= NMR SPECTRUM
##DATA CLASS= XYDATA
##OBSERVE FREQUENCY= 399.682468187609
##OBSERVE NUCLEUS= ^1H
##XUNITS= HZ
##UNITS= ARBITRARY UNITS
##XFACTOR= 0.250974584506556
##YFACTOR= 1
##DELTA= -0.250974584506556
##NPOINTS= 32768
##IRSTYX= 6579.37661293139
##FIRSTY= 0
##LASTY= -1644.30759759492
##MAXY= 383963588
##MINY= -198600
##XYDATA= (X+1(Y, Y))
32767a0140m2061L521J1389L6231535r502m659J3512K621o270q6
32753G519K076j625Q89o742j67m861P042o661L73N675k719L534
...
Oa112
##END=
  
```

1. Identification of the signals
 2. Integration and multiplicity determination
 3. Assignment of each signal to the corresponding atom in the molecule
 4. Exportation of the data for publication and/or for database storage

1. Just test the applet with ethylmethylacetate or upload a fully assigned spectrum, click on the spectrum and check the results over the peak labels.
 2. Assign the spectrum with the "Smart Peak Picking" tool.
 3. Draw a molecule, test with ethylmethylacetate or upload a molecule.
 4. Get the result in ACS format or as a table of assignments.

¹H NMR δ 6.47 (dd, J = 14.4, 7.0 Hz, 1 H), 4.20 (dd, J = 14.4, 2.1 Hz, 1 H), 4.00 (dd, J = 6.7, 1.8 Hz, 1 H), 3.77 (q, J = 7.0 Hz, 2 H), 1.31 (t, J = 7.0 Hz, 3 H)

Fig. 1. After acquisition on the NMR spectrometer, the spectrum is processed (i.e. on Bruker TopSpin: ft, apk, abs) and the real/imaginary parts exported to a jcamp-dx text file (i.e. on Bruker TopSpin: tojdx, Output data: SPEC REAL + IMAG, Compression mode: DIFF/DUP). This jcamp-dx file can be uploaded in the 'NMR Assigner' on the web site <http://www.nmrdb.org> (1). A chemical structure can either be drawn directly from the site or uploaded as a molfile (2). Assignment of the peaks is done by drawing a horizontal line around the different signals using the 'Smart Peak Picking' tool. Simultaneously, the program will integrate the signal and try to extract the multiplicity, chemical shift and coupling constants. In case of overlapping signals, massifs can be defined or peaks composing a multiplet can be selected individually (3). Complete assignment of the molecule is done by drawing lines between the labels and the corresponding atoms (4). Finally exportation for publication or for future processing is done by clicking on the corresponding link (5).

for publication or as tab-delimited text for further processing such as the storage in a database.

NMR Resurrector: Resurrecting NMR Spectra

Millions of NMR spectra have been published in the literature during the last 40 years. However, most of this information is in the form of string descriptions, hardly the ideal format for complex searches and unsuitable for efficient human visualization for all but the most simple molecules. In order to recover this information, we have developed a tool that analyses NMR

string assignment and recreates a fully-annotated, visual and interactive representation of the original spectrum (Fig. 2). Once this is achieved, the resurrected data can be handled like any other spectra, for example as a reference in the comparison with an acquired spectrum or as source for the correct storage in the database.

Conclusions

These first tools show what is today possible to do from the web and pave the way towards a virtual lab journal where all the information from the laboratory is stored in databases and can be added and processed

1. Copy and paste the NMR assignment string.
 2. Click here to resurrect the spectrum. You may now click on the spectrum and check each assignment by going over the peak labels.
 3. Draw a molecule, test with this phosphonate derivative or upload a molfile.
 4. Get the result in ACS format or as a table of assignments.

Fig. 2. Copy/paste of the NMR description of a NMR spectrum found in the literature (1) followed by a click (2) allows a NMR spectrum to be reconstructed. A mouse over the Smart Peak Labels will show in the bottom of the applet the specific information about the signal (3).

from any web-browser on any operating system. This allows not only to keep all the information safe but also to guarantee future retrieval of the information.

An extension of these tools allowing submitting and processing directly spectra in a central database is currently in development.

Acknowledgements

We thank Agfa-Gevaert and Union Chimique Belge for financial support, Thomas Sander from Actelion for structure depiction code.

Received: February 25, 2008

- [1] R. S. McDonald, P. A. Wilks, *Appl. Spectrosc.* **1988**, *42*, 151.
- [2] P. Lampen, H. Hillig, A. N. Davies, M. Linscheid, *Appl. Spectrosc.* **1994**, *48*, 1545.
- [3] A. N. Davies, P. Lampen, *Appl. Spectrosc.* **1993**, *47*, 1093.
- [4] P. Lampen, J. Lambert, R. J. Lancashire, R. S. McDonald, P. S. McIntyre, D. N. Rutledge, T. Frohlich, A. N. Davies, *Pure Appl. Chem.* **1999**, *71*, 1549.
- [5] TopSpin, <http://www.bruker-biospin.com/topspin.html>.
- [6] C. Steinbeck, S. Krause, S. Kuhn, *J. Chem. Inf. Comput.* **2003**, *43*, 1733.
- [7] L. Patiny, *Internet J. Chem.* **2000**, *3*, n°2.
- [8] Spectral Database for Organic Compounds, SDBS, <http://riodb01.ibase.aist.go.jp/sdb/s/>.
- [9] L. D. Paulson, *Computer* **2007**, *40*, 24.
- [10] *ACD/SpecManager*, <http://www.acdlabs.com/>.
- [11] J. C. Cobas, F. J. Sardina, *Concepts Magn. Reson. A* **2003**, *19A*, 80.
- [12] G. Balacco, *J. Chem. Inf. Comput.* **1994**, *34*, 1235.
- [13] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277.