Research article

MRC

Received: 20 February 2015 Revised: 10 April 2015 Accepted: 20 April 2015 Published online in Wiley Online Library: 21 May 2015

(wileyonlinelibrary.com) DOI 10.1002/mrc.4263

Facilitating quality control for spectra assignments of small organic molecules: nmrshiftdb2 – a free in-house NMR database with integrated LIMS for academic service laboratories

Stefan Kuhn[a]* and Nils E. Schlörer[b]**

ABSTRACT: nmrshiftdb2 supports with its laboratory information management system the integration of an electronic lab administration and management into academic NMR facilities. Also, it offers the setup of a local database, while full access to nmrshiftdb2's World Wide Web database is granted. This freely available system allows on the one hand the submission of orders for measurement, transfers recorded data automatically or manually, and enables download of spectra via web interface, as well as the integrated access to prediction, search, and assignment tools of the NMR database for lab users. On the other hand, for the staff and lab administration, flow of all orders can be supervised; administrative tools also include user and hardware management, a statistic functionality for accounting purposes, and a 'QuickCheck' function for assignment control, to facilitate quality control of assignments submitted to the (local) database. Laboratory information management system and database are based on a web interface as front end and are therefore independent of the operating system in use. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: NMR; lab administration system; LIMS; NMR database; assignment quality control; prediction; structure verification; nmrshiftdb2

Introduction

With the increasing availability of modern medium and high-field NMR spectrometers for routine characterization of small molecules, the 'output' of data from academic institutions has grown considerably during the last two decades. However, while state-of-the-art equipment is available to generate the data, solutions for the post-processing steps like data treatment, assignment of spectra, and quality control of NMR data submitted for publication often lag behind. One freely available aid for academic service laboratories is offered by the local database and assignment tools of nmrshiftdb2 and its integrated laboratory information management system (LIMS), which can facilitate the handling of NMR spectra on the downstream side of the NMR facility.

Background

Academic NMR service facilities nowadays often need to face two main issues: on the one hand, a low budget and a reduced personnel situation and, on the other hand, an exponential growth in (NMR) data. The latter is due to the increasing availability of modern NMR equipment (including robots and automation) and pulse sequences, allowing to run experiments for an 'over-determination' of structure assignments in a fraction of the time that was necessary in the past. Thus, although the efficiency of NMR for accurate assignment has evolved rapidly after the introduction and availability of gradient-selected 2D experiments in academic service laboratories, the manner how chemistry research labs/groups deal with the resulting data often does not keep pace with the increasing amount of 'assignable' structures. This becomes manifest in an observable decline of assignment quality.[1] In fact, several journals for synthetic chemistry do not require full assignment or actually reduce data presentation, and there is an increasing number of cases of 'data mismanagement' in Supporting Info sections,[2] not to mention the reduced quality of assignments, including incomplete and wrong assignment, leading in worst case to later retraction of data.

At the same time, in a field of such paramount importance for day-to-day structure elucidation like NMR spectroscopy, there still exists the following: (i) no agreement on how to generally submit experimental data for small molecules[3] and (ii) there are no (freely available) protocols/tools/implementations for a quality check of assignments for NMR data of small molecules, as it is in contrast the case (and often required), e.g. for the submission of X-ray and

* Correspondence to: Stefan Kuhn, Department of Computer Science, University of Leicester, Leicester, UK. E-mail: shk12@le.ac.uk

** Correspondence to: Nils E. Schlörer, Department für Chemie, NMR-Abteilung, Universität zu Köln, Greinstr.4, 50939 Köln, Germany. E-mail: nils.schloerer@uni-koeln.de

a Department of Computer Science, University of Leicester, Leicester, UK

b Department für Chemie, NMR-Abteilung, Universität zu Köln, Greinstr.4, 50939 Köln, Germany

peptide/protein NMR data.[4] In addition, the occupation of NMR facility staff, with an increasingly reduced number of positions available, often does no longer allow to implement 'traditional' ways of spectra assignment and quality control of assigned spectra for most samples. Thus, the consequent treatment of scientific data is often disregarded or left to the attention of scientists not sufficiently trained in this specific field, because of the lacking availability of the aforementioned resources. Therefore, in spite of the educational efforts to teach the information contents of NMR spectra *per se*, users are often 'left alone' with the processing and assignment procedures for their (high quality) data. Additionally, they are often not informed about the possibilities on how to evaluate their assignments (besides using the 'predict function' of commercially available chemistry software, which can be a limiting factor for academic scientists). Databases for small organic molecules, which make available free tools to check the quality of assignment, are scarce[5] or focused on specific classes of compounds.[6] Large repositories are either restricted to internal use in industrial laboratories or only commercially available.[7]

Recommendations from scientific research councils or scientific divisions for magnetic resonance[8] include the entire number of steps from making measured data available for users, local data storage, and data handling. However, so far, these recommendations have mainly been imposed on 'large-molecule' NMR laboratories, where throughput and workflow are comparatively minor limiting factors. For day-to-day, small-molecule NMR facilities in academia, which have to serve many different user groups at the same time, the practical situation is often significantly more complex. One aspect to improve this situation to a certain extent would be to locally improve the assignment quality 'on the roots', i.e. provide service laboratories and users with an infrastructure that allows a more efficient data treatment for assignment and its quality control. This was the origin of our design for a LIMS for an academic service NMR laboratory, which we describe in the present paper (Fig. 1).

With the intention to install a free, open NMR database for small molecules, NMRShiftDB was created in the early 2000s.[9] This database is so far the largest, completely free installation of its kind. To integrate NMR users and to make them familiar with the handling of the web-based software and the integrated tools, during the last 8 years, we programmed, installed, and institutionalized a LIMS for nmrshiftdb2, at the University of Cologne (UzK), serving as an interactive, ongoing project to improve data quality control of NMR assignments. At this point, it features most of the aforementioned requirements, is platform-independent, is easy to implement, and has reached a stage of functionality and documentation, where we would like to introduce it to the NMR community and invite others to participate, share, use, and develop it as a common tool for academic service laboratories.

Experimental and methods

Local situation – NMR facility at the University of Cologne

At the Department of Chemistry of UzK, NMR spectrometers are operated partially in open access (four instruments), but also by technicians that belong to NMR staff (three instruments). The facility is supervised by a scientific manager. Besides an introduction to NMR spectroscopy, 1D and 2D NMR spectra assignment is taught

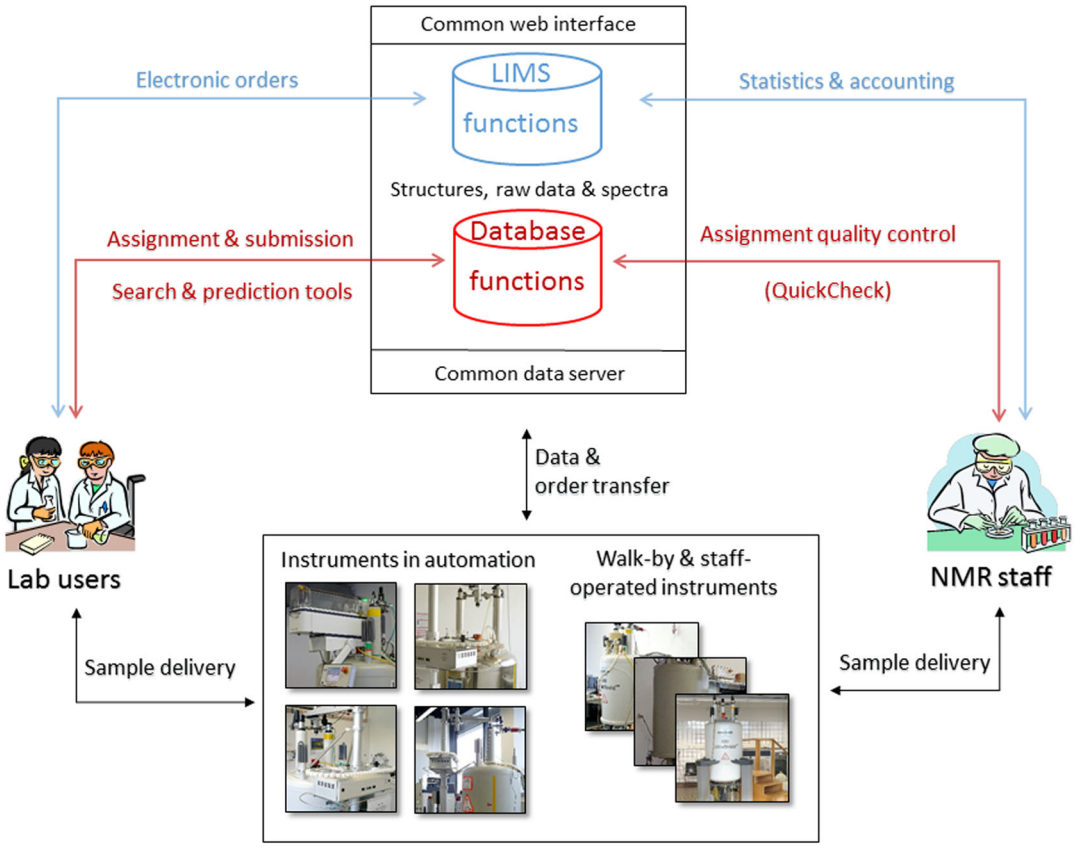

Figure 1. Functions of nmrshiftdb2's laboratory information management system (LIMS) and database system in a lab environment. A common server operates the lab administration functions (blue) and provides the functionalities of a local database with assignment tools and (optional) access to the World Wide Web database of nmrshiftdb2 (red).

as an essential part of chemistry education; also, courses on the use of NMR databases and, more specifically, the tools provided by nmrshiftdb2 are obligatory. The service facility is an integral part of the medium-sized department with about 150 active NMR users from organic, inorganic, physical and biochemistry, and some additional external academic institutions. Users include all levels of chemical education, ranging from undergrad synthetic chemistry laboratory classes through advanced projects to students conducting their thesis and research fellows. Before being able to make use of the NMR facilities, each user is required to register and receives a short introduction on nmrshiftdb2's LIMS. Currently, about 35 000 samples are processed per year. Although in our department, instruments manufactured by Bruker (Bruker Biospin GmbH, Rheinstetten, Germany) are operated and the corresponding software (TopSpin vs 1.3–3.2 and ICON-NMR vs 4.0–4.6)[10] is used, nmrshiftdb2 can be applied to other environments, as well, because it is platform independent. An additional installation is currently under way at the NMR facility of Ruhr University Bochum (RUB, Germany) and the NMR lab of the Institute of Organic Chemistry at Johannes Gutenberg University (Mainz, Germany).

nmrshiftdb2 – LIMS and local database

nmrshiftdb2 provides two basic functions, a LIMS and a local database: The LIMS operates the data server, processes orders for automated measurements, and allows access to orders through two different profiles, i.e. facility users and facility staff. It includes tools for a fully electronic lab administration – a 'paper-free' lab. For this, we link the nmrshiftdb2 system to the spectrometers in two ways. On the one hand, orders must be transferred to the instrument manufacturer's software if a sample changer is used. In the current setup, this is carried out by copying text files that comply with the Bruker ICON-NMR format for 'External Setup' via Secure Copy (SCP) to the TopSpin computers.[10] On the other hand, files produced by spectrometers must be transferred into the nmrshiftdb2 system. For this, we use a script that automatically transfers the files generated by the spectrometers on the nmrshiftdb2 server using the Server Message Block (SMB) protocol. This setup can be varied depending on the environment in the lab and can also be adopted for other instrument manufacturers. Currently, the free scheduling service 'faces'[11] is integrated for scheduling purposes. The orders are used to produce statistics, based on the types and numbers of experiments carried out by the members of the groups, other departments, and external users. For accounting taking instrument times into account, a closer integration of the scheduling function is envisaged.

The local database function allows to build a dedicated database, which can be used similar to the public nmrshiftdb2 instance, but can contain local data in addition to data from the public server. These data may be confidential and will help get better predictions specifically in the areas local groups are working on. Similar to the public database, users can submit data, but they can do this from their sample orders, using existing data like probable structure. They will also see a prediction and be able to use the CSEARCH robot referee.[5b,12] This links the lab system with the quality control. At the same time, administration of the NMR facility becomes easier and contributes to the overall goal of quality control.

Modes of installation, practical implementation

The software must be installed on a server, which is accessed by the users via a web browser. The hardware requirements are modest; any current ×86 server is sufficient. Technologies like high-availability systems or Redundant Array of Independent Disks (RAID) disks can be used as wished. Because the software is written in Java, it can run on any operating system where a J2EE-compatible servlet container can run, which should be all current systems including Linux and Windows. Detailed installation instructions are contained in the nmrshiftdb2 download.[13]

The installation at the UzK includes, besides the LIMS, a local database, with a protocol to assure privacy of data (and later submit to nmrshiftdb2's World Wide Web database). However, it should be mentioned that LIMS and local database also can be installed separately: A lab that runs its instruments differently can still build a local database. For privacy reasons, such an instance can be operated without being visible from outside. Users can also have private entries not visible even to other internal users. In order to profit from updates on the public database, a one-way replication from the public to the local database can be set up. Further details (explicit instructions for the use of nmrshiftdb2's LIMS and descriptions of the functionalities and use of the database) are provided in the Supporting Information available for this article.

Results and discussion

In the following, after a short section on the installation and development of nmrshiftdb2, we describe in more detail its functionalities and how LIMS and local database installations can be adopted to environments different from the ones where it is currently running.

The current version of nmrshiftdb2 has been developed from the original database NMRShiftDB of C. Steinbeck and S. Kuhn[9] and included from the beginning as a central component an electronic order submission and processing functionality. In collaboration with the chemistry students of the department's organic chemistry research groups, functionality and design of user surfaces have experienced continuous development and revision.

Organizing the lab using LIMS: sample submission, processing, and access to data

The LIMS supports various modes for the administration of a service facility. However, to give an impression of the mode of operation, we illustrate the functionalities for the environment at UzK. Here, a combination of hands-on, robot-supported, and staff-run spectrometers might be representative for most of the prevailing types of access for service instruments.

Currently, the installation of nmrshiftdb2's LIMS at UzK is based on three different user profiles:

- Students and scientific users who make use of the NMR services can interactively submit samples and electronic orders and download raw data of completed experiments, with the option for a subsequent (software-assisted) assignment and/or submission of assigned data for review to the local database.
- A second profile exists for NMR staff members, who can view incoming orders for 'manual processing' on research spectrometers (which, at UzK, are not accessible for hands-on measurements), as well as all electronic orders submitted by users for instruments running in automation mode (if desired).
- Finally, an administrator account allows the NMR laboratory manager to oversee, in addition to the aforementioned staff, functionalities to define spectrometers, experiments, and modes of operation for order forms; grant access to and/or close user accounts; and accept or refuse assigned datasets

submitted to the local database. An important function available in the latter type of profile is the possibility to generate user statistics for accounting. This is currently performed using the types of experiments per user and group. Also, instrument types are distinguished. So far, time consumption is not taken into account.

So far, our LIMS has been tested with the following modes of spectrometer operation:

- As mentioned before, open-access spectrometers in automation or routine mode can be handled.[14] Data are transferred from the spectrometer to the local server via the Server Message Block (SMB) protocol. On Bruker software, a script can be integrated to do this; similar setups are possible with other manufacturers. Once data directories have arrived at the server, nmrshiftdb2 attaches zipped raw data to the corresponding order forms and makes them accessible for download by users through the common web interface.
- Also, operator-run spectrometers, which are supervised by NMR staff, are integrated in a similar manner, allowing operators to choose at which stage they want to transfer the data to the local server, where the order then will be automatically completed, as well.
- If necessary, the LIMS also allows to attach data 'manually' from the server by the users themselves, in case that such is required for local conventions – while the automated attachment through the software might have a duration of a maximum 60 min (can be configured in the installation); researchers that are used to get their data directly from the server can remain with their habits; however, each data transfer between spectrometer and nmrshiftdb2 server will be accounted for the statistics of the lab (Figs 2–4).

In order to show how LIMS and database work in practice, we describe the steps from the submission of an order for a spectrometer operating in automation to a finished entry in the database.

- The user connects to the local nmrshiftdb2 instance and logs in with his or her account.
- He or she fills in the electronic order form and chooses the desired experiments (e. g. ^{13}C, ^{1}H standard spectra and 2D correlations). The user also submits a probable structure[15] and other details, including the holder number where he or she leaves his sample.
- The order is automatically entered into the spectrometer software and executed. The spectra files generated are picked up by nmrshiftdb2 software.
- The user logs in and can see if the order has been finished. He or she can download the resulting files and analyze them using processing software.
- Once the user is confident about the structure, he or she can switch from the 'NMR lab administration' page to the 'submit' page. Data including the probable structure are already filled in. The user can either enter the peaks manually or perform a peak picking on a JCAMP-DX file. Using the semi-automated assignment and comparing this with the assignment yielded from experiments, he or she submits the assigned peak spectrum.
- The assignment is automatically submitted to the CSEARCH robot referee, whose rating and evaluation report are saved alongside the assignment. If the CSEARCH rating is 'accept' or 'minor changes', the lab administrator (who might be identical with an NMR staff) will review the assignment.
- The record is at this point also ready to be exported into publication formats. It is available via a fixed URL, including the quality report and all meta-information that were entered when it was submitted for review.

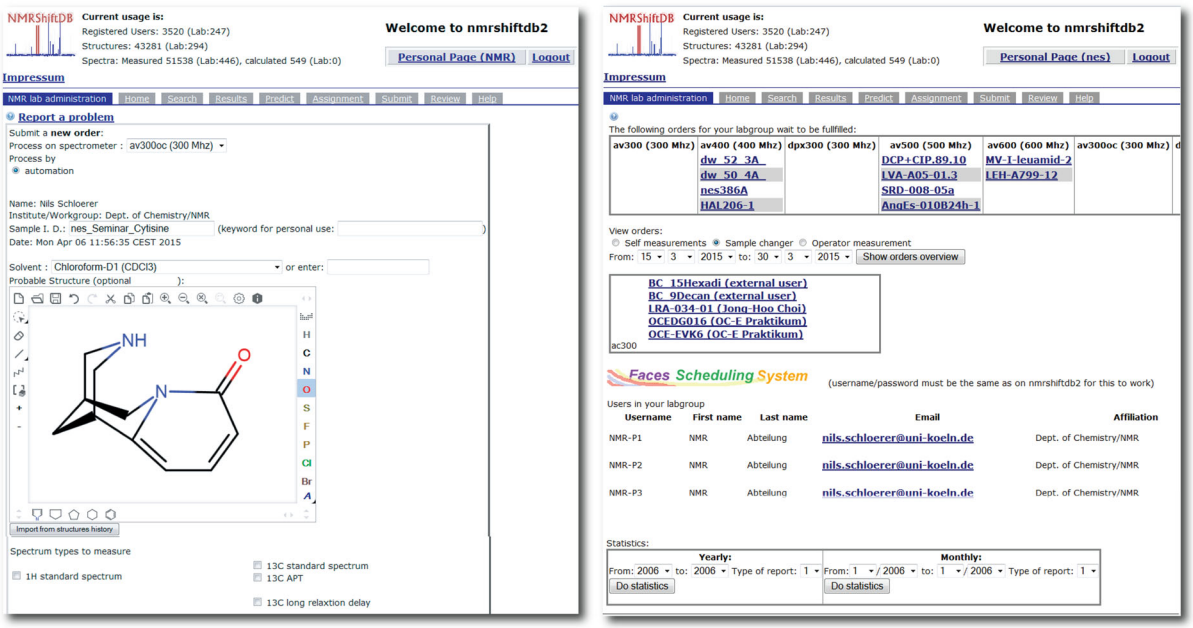

Figure 2. The laboratory information management system of nmrshiftdb2: details from the order form for users and the staff interface. The order form (left) allows the selection of experimental conditions – spectrometer, solvent, and experiment – and the mode of order processing – automation, hands-on, and operator. If desired by the lab administration, entering a structure may be mandatory. The individual settings can be defined by an NMR staff on the staff interface (right). Through the latter, administration and order processing can be supervised, which allows, amongst other options, to see and complete open orders by staff, to oversee the workflow, to schedule measurement time, or to generate statistics for accounting.

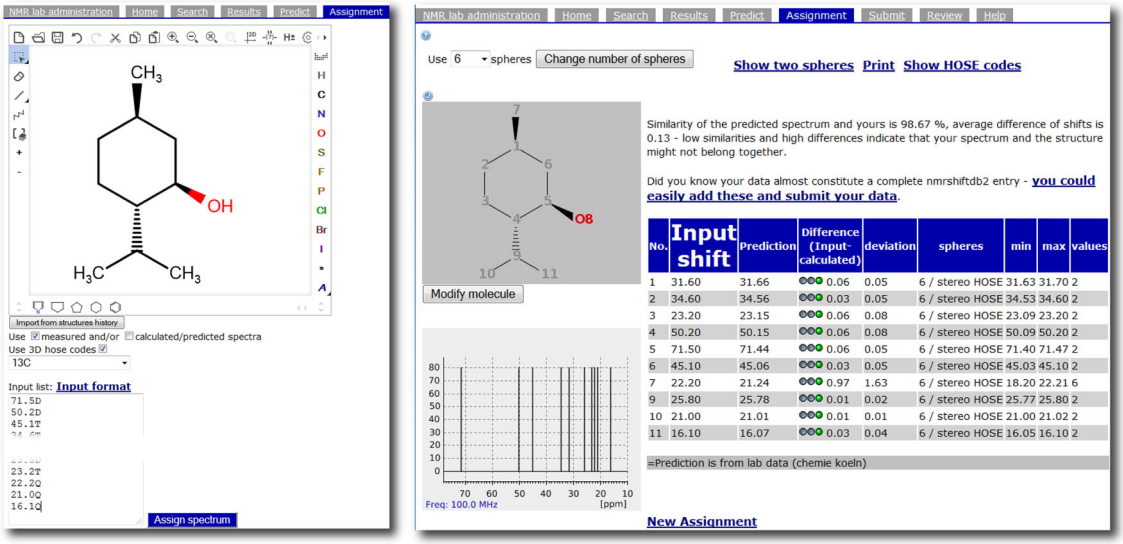

Figure 3. Using the assignment aid of nmrshiftdb2. After submission of a structure and a peak list (left), the results of an automated assignment are presented with a color-coded evaluation (right). Green on chemical shifts indicates a close fit of the assignment with the predicted spectrum. Atoms, where assigned and predicted chemical shifts show a larger divergence, appear in yellow or red, indicating that no match between structure and spectrum can be built.

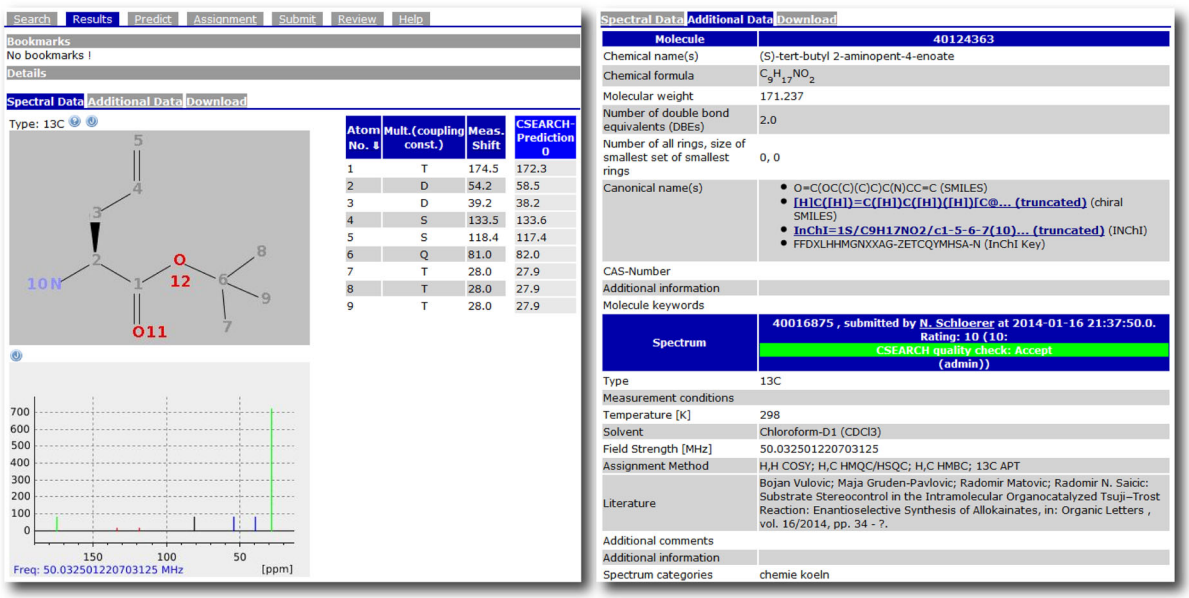

Figure 4. Evaluation of an assignment by nmrshiftdb2's QuickCheck: Quality of the assignment as judged by CSEARCH's[12] RobotReferee and nmrshiftdb2's prediction is presented by a color-coded report and an evaluation rating from 1 to 10. In addition, a quality report by CSEARCH can be accessed.

This process allows users to submit paper-free orders, helps the lab with keeping electronic records, and offers the resulting files to users, ready for download. At the same time, it assists users with assignments and NMR staff with quality assurance.

Assignment tools and local database: improving quality of published data

The local database installed at UzK comes with the same range of functionalities, which are available for our free, open web-based database at www.nmrshiftdb.org: Access is granted to use the more than 40 000 structures, which have been deposited so far, for extensive search and prediction options. In addition, locally entered assigned structures will be included in the functionality, thus providing a better performance once enough fragments characteristic for the local research lines have been stored.

For several reasons, nmrshiftdb2 should be considered as a versatile, supporting tool for NMR-based structural analysis:

- It contains a prediction function for various NMR-active nuclei (so far, mostly ^{1}H and ^{13}C chemical shifts are populated; however, for local databases, any other isotopes can be included).

Thus, expected chemical shifts for a spectrum can be yielded, based on an improved Hierarchical Organisation of Spherical Environments (HOSE) code algorithm,[16] which takes stereo configurations into account. Shifts with up to six spheres and identical stereo configuration are searched; fallback to less spheres and disregarding the stereo configuration is then applied. The user is warned if values retrieved are unlikely to be reliable; thus, no false impression of reliability is given. HOSE code prediction with four or more spheres and respecting stereochemistry are generally considered reliable. Especially if in-house data are included, this can be achieved for the vast majority of predictions.

- The database offers various options to perform searches, e.g. for the dereplication of compounds, which are known. For such purpose, either some or all chemical shifts can be entered as search key. Instead, also, the reverse approach can be chosen: By means of a (sub)structure or similarity search, the range of chemical shifts available in the database can be investigated.
- Besides, a semi-automated assignment tool can be used for guiding users through the structure elucidation process: If a (peak) list of chemical shifts and a suggested structure are available, nmrshiftdb2 provides an assignment based on its prediction algorithms. The user can see the quality of the prediction, as explained before, and how well his or her shifts match the prediction in the suggested assignment. The user can either accept the assignment or edit it. Ultimately, it can be exported as a text section for 'copy and paste' into reports or publications.

A most recent development is a 'Quick Check' protocol, which consists of the submission of the input to the build-in evaluation according to nmrshiftdb2's prediction. The latter is available immediately during submission and is presented together with a confidence interval for each shift value. In case of ^{13}C chemical shifts, the data are also submitted to the CSEARCH robot referee. Entries, which yield an evaluation of the level 'accept' or 'minor revisions' there, are passed on to NMR staffs, which do an expert check. Entries accepted by them are considered approved and ready for publishing. Approved datasets were checked by a combination of nmrshiftdb2's prediction, CSEARCH prediction (for ^{13}C data), and NMR staffs. This helps raise the quality of assignments that are published, and relieves NMR personnel from going through obviously wrong entries.

Although the 'Quick Check' process includes several steps, benefits outweigh the necessary efforts for users: Users receive the nmrshiftdb2 prediction and get the CSEARCH assessment in case of ^{13}C NMR spectra; a permanent, electronically accessible database record is created, and the assignment can be easily exported to various formats. In a local installation, the users primarily contribute to a local database, thereby improving predictions for compounds typically used in the local laboratories, but at that point not available in the public database. Transfer of data to the public instance of nmrshiftdb2 either might be individually granted by the lab manager or can be automatized: Publication is restricted to structures already present in PubChem, which ensures no intellectual property is released to the public inadvertently. If the option to submit original data with the assignment is used, an electronic archive for coworkers is built, which is especially useful if the author has left the institution.

Implementation, training, and practical aspects

The implementation of a local database, integrated with a LIMS, is desirable for improving research, as well as NMR laboratory management. However, the changes brought about for the accessing user groups and the institution of which the NMR lab forms a part will most often require to take some accompanying educational and administrative measures, which will be shortly considered here.

In terms of hardware and programming, efforts connected with the setup of the LIMS and the database will be low. Also, on the side of the staff, only the initial configuration of lab administration parameters like spectrometers, the available modes of operation, and the types of experiments requires additional setup time. On the training side for users, the installation should be accompanied by explanatory introductions and seminars or workshops, intended to grow a sense of responsibility for the issue of assignment and data quality. These included in our case workshops for users, as well as the active participation of users in development: To facilitate the adaptation to local requirements, we integrated (selected) local users for 'beta testing'. This provided direct feedback and served as vehicle to increase the acceptance of the LIMS for users. Integrating users as early as possible in the implementation of the LIMS, when accompanied by explanation on the scientific background, improves the understanding and valuation of assignment quality control, going beyond the purely administrative use of the system. At UzK, introduction of users is divided into different levels:

- A short, general introduction to the use of the LIMS is given for freshman students when entering the first lab using NMR spectrometers in automation, including a hands-on training on the spectrometer by NMR staff for small groups of four to five students. Thus, statistics for the use of spectrometers by teaching labs are handled through corresponding lab accounts that can be conveniently installed and serve as a general user account for all students of the teaching lab.
- During a later stage of their chemistry studies, in an interactive training course, the importance of a database for assignment quality and the use of its functionalities (by means of a quiz) are introduced to MSc students in our computer lab.
- In addition, students who enter a research group to perform independent scientific projects (in our case at MSc level) are provided with their own account, which is granted to them after a short quiz including the regulations and use of the facility (groups of three students).

The advantages of the nmrshiftdb2 lab system consist in a fully electronic (i.e. paper-free) administration, giving transparency to users and the lab, which facilitates accounting and statistics. As a main added scientific value, the uniform user interface for lab administration and the database and assignment tools need to be emphasized, which encourage users to get involved with the database and to develop a sense for assignment quality. Compared with a completely non-supervised system with free access to either each spectrometer workstation or an open, central spectra server on a self-service basis, even basic tasks like setting up accounts for new users and enforcing an electronic administration system are initially connected with training investments.

At UzK's chemistry department, the nmrshiftdb2 lab system has been running since 2007, and all orders are processed via the LIMS. Self-service and operator service are used. For automation mode, the orders are transferred to Bruker software, making nmrshiftdb2 the single point of interaction for users. Assignment by users and database usage has been started, although this is an ongoing process. Several aspects are still under development. For example, a closer integrated scheduling system would be desirable, making accounting

based on time usage possible. The system is also in constant adaptation with respect to database features like the prediction system.

Conclusion

A fully electronic lab process offers advantages for users and facility managers/staff. Users can see the current state of their orders and track their data. Additionally, integration with the database functions makes it easier to familiarize users with proper data treatment and advanced database techniques, which ultimately improves the quality of the published data. The CSEARCH robot referee, serving as an aid for the evaluation of semi-automated assignment, is a key functionality for this. On the other hand, NMR staff enjoys full transparency and can compile statistics for accounting purposes 'with a keystroke'. With the described implementation, even with the comparatively large number of data produced in our service facility, a quality control for assignments for data to be published in thesis and other scientific reports becomes feasible even with a low staff/spectra ratio. Also, (ab) use of expensive spectrometer time and the distribution of samples to useful applications can be better supervised. As intended in our department, the NMR facility operates paper-free.

The advantages come at a price: Compared with a completely non-supervised system with free access to either each spectrometer workstation or an open, central spectra server on a self-service basis, even the basic tasks of setting up accounts for new users and enforcing an electronic administration system are initially connected with educational investments and need support or at least acceptance from the scientific institution.

Currently, the database is restricted to 1D NMR experiments, whereas the LIMS can be used to 'order' any NMR experiments. An extension of the database to 2D experiments, also in prediction and assignment, is a step currently under development. Although scalar couplings (homonuclear and heteronuclear) can be included during the electronic assignment, at the current state of development, they are not yet used as an integral part in interactive structure elucidation. A high priority for the LIMS is the inclusion of a native scheduling system and a flexible accounting facility. Also, integration of a barcode reader system is currently under testing. Both, LIMS and database, could ultimately also include other techniques (MS, IR, UV/Vis, etc.) because of the modular character of setup, for LIMS, as well as for database.

We are convinced that other departments and laboratories can profit from the currently available, free software and our experiences made so far. Because of the open-source nature and the flexibility of the software, it should be possible to implement it in different settings. Our desire is to support nmrshiftdb2 as a free, openly accessible database with LIMS and to encourage its integration in the NMR community. Therefore, we invite interested parties to collaborate with us on the further development of this free lab system and, thereby, on a better data and assignment quality in the NMR sector.

Acknowledgements

We would like to thank Dr Erhard T. K. Haupt (Universität Hamburg), Dr Matthias Jank (Bayer Crop Science), Heinz Kolshorn (Universität Mainz), Dr Johannes Liermann (Universität Mainz), Prof. Dr Wolfgang Robien (Universität Wien), Prof. Dr Raphael Stoll (Ruhr Universität Bochum), and Dr Christoph Steinbeck (EMBL-EBI) for their helpful discussions and support.

An important point for any free and open access development is institutional support and acceptance. In a university setting, it is particularly the principal investigators and the research group leaders, whose backup decides upon the successful introduction. We are therefore indebted to the members of the Department of Chemistry, University of Cologne, and the regional center for computation (RRZK), University of Cologne, for the patronage during the development and implementation of our LIMS.

References

[1] W. Robien. *Trends Anal. Chem.* **2009**, *28*(7), 914–922.
[2] a) For some recent discussions, see: B. Halford. *C&EN* **2014**, *92*(16), 32–33; b) J. A. Gladysz, R. B. Bedford, M. Fujita, F. P. Gabbaı, K. I. Goldberg, P. L. Holland, J. L. Kiplinger, M. J. Krische, J. Louie, C. C. Lu, J. R. Norton, M. A. Petrukhina, T. Ren, S. S. Stahl, T. D. Tilley, C. E. Webster, M. C. White, G. T. Whiteker. *Organometallics* **2014**, *33*, 1505–1527; c) A. B. Smith III. *Org.Lett.* **2013**, *15*(12), 2893–2894; d) E. E. Podlesny, M. C. Kozlowski. *J.Nat.Prod.* **2012**, *75*, 1125–1129; e) M. Elyashberg, A. J. Williams, K. Blinov. *Nat. Prod. Rep.* **2010**, *27*, 1296–1328.
[3] M. J. Harvey, N. J. Mason, H. S. Rzepa. *J. Chem. Inf. Model.* **2014**, *54*, 2627–2635.
[4] a) A repository for data from NMR spectroscopy on proteins, peptides, nucleic acids, and other biomolecules: *Biological Magnetic Resonance Databank* (*BMRB*). http://www.bmrb.wisc.edu/ [18 February 2015]; b) A repository for small-molecule organic and metal-organic crystal structures: *Cambridge Structural Database* (*CSD*). http://www.ccdc.cam.ac.uk/Solutions/CSDSystem/Pages/CSD.aspx [18 February 2015]; c) The development of the RSC open access portal ChemSpider is described in: A. Williams, V. Tkachenko. *J. Computer-Aided Mol. Design* **2014**, *28*(10), 1023–1030; d) For a recommended list of repositories by *nature*, see http://www.nature.com/sdata/data-policies/repositories#molec [18 February 2015].
[5] a) A freely searchable database for infrared, NMR and mass spectra of organic molecules: *integrated spectral data base system for organic compounds* (*SDBS*). http://sdbs.db.aist.go.jp/sdbs/cgi-bin/cre_index.cgi [18 February 2015]; b) A free tool to evaluate the quality of the assignment of ^{13}C NMR data: *CSEARCH robot referee*. http://nmrpredict.orc.univie.ac.at/c13robot/robot.php [18 February 2015]; c) A free NMR database for organic molecules and their NMR spectra (^{13}C, ^{1}H and other nuclei): *nmrshiftdb2*. http://www.nmrshiftdb.org [18 February 2015].
[6] Since there is a wealth of NMR databases, which are specialized on particular aspects of different fields of molecular research, we decided not to mention any of them for lack of space.
[7] a) A library of high resolution NMR spectra (^{1}H, ^{13}C and other): *ACD/labs databases*. http://www.acdlabs.com/products/dbs/nmr_db/ [18 February 2015]; b) Commercial database for ^{13}C chemical shifts, partially included in Bio-Rad's database, NMRpredict or SpecInfo: *CSEARCH*. http://csearch-nmr-data.blogspot.de [18 February 2015]; c) Commercial NMR database (various nuclei): *BioRAD's HaveItAll*. http://www.bio-rad.com/de-de/product/spectral-identification/nmr-databases?pcp_loc=catprod [18 February 2015]; d) Commercial database used for the prediction of chemical shifts and NMR spectra: *NMRpredict*. http://www.modgraph.co.uk/product_nmr.htm [18 February 2015]; e) Commercial spectra database: *SpecInfo*. http://www.wiley-vch.de/stmdata/specinfo.php. DOI: 10.1002/9780471692294 [18 February 2015].
[8] a) Representative for such recent developments, the recommendations of the European Research Council (ERC) and the German Research Council (DFG) for scientific data treatment and core facility organization might serve as examples: Basic requirements for research infrastructures in Europe by ERC (March 2011), http://www.dfg.de/download/pdf/foerderung/programme/wgi/basic_requirements_research_infrastructures.pdf [18February2015]; b) Requirements for core facilities by the DFG, http://dfg.de/en/research_funding/programmes/infrastructure/scientific_instrumentation/core_facilities/index.html [18 February 2015].
[9] a) C. Steinbeck, S. Kuhn. *Phytochemistry* **2004**, *65*(19), 2711–2717; b) C. Steinbeck, S. Krause, S. Kuhn. *J. Chem. Inf. Comp. Sci.* **2003**, *43*(6), 1733–1739.
[10] *TopSpin and ICON-NMR*, Bruker Biospin GmbH, Rheinstetten, Germany.
[11] W. York, S. Roushanzamir, Complex Carbohydrate Research Center: Faces Scheduling System (Version 3.1) [Software], **2003**. Available from http://faces.ccrc.uga.edu [18 February 2015].
[12] a) For references on CSEARCH, see: L. Chen, W. Robien. *Chemometr. Intell. Lab.* **1993**, *19*, 217–223; b) H. Kalchhauser, W. Robien. *J. Chem.*

Inf. Comp. Sci. **1985**, *25*, 103–108; c) http://homepage.univie.ac.at/Wolfgang.Robien/csearch_main.html [18 February 2015].
[13] Installation guide and links, as well as detailed information can be found in the Supporting Materials section.
[14] So far, Bruker Biospin's TopSpin software with the corresponding automation software packages of ICON-NMR (Versions 4.0-4.6) has been used. However, an implementation for Agilent or Jeol spectrometers will not present major difficulties, as far as setup through text files can be established.
[15] Marvin JS is used for drawing and displaying chemical structures, Marvin JS 14.10.27, **2014**, ChemAxon (http://www.chemaxon.com).
[16] W. Bremser. *Anal. Chim. Acta* **1978**, *103*, 355–365.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site.