Manual for the spectra database functions of nmrshiftdb2:¹
(Revision 15-12-2014):

This is a “how to” with instructions for the use of nmrshiftdb2 (predict, search and assignment aid/submit functions). Nmrshiftdb2 is a database program, accessible via its web surface at http://www.nmrshiftdb.org (www installation at Cologne University). For a more detailed description than the one given below, please check out our Help pages at the nmrshiftdb2 homepage.

1.1 Predicting chemical shifts of a molecule (¹³C, ¹H)

Nmrshiftdb2 offers the prediction of spectra for all nuclei. However, the quality of the prediction depends (amongst other) on the number of spectra/structures for the corresponding nucleus type stored in the database. I.e., for “exotic” nuclei like ²⁹Si, you will get a prediction, but its value is questionable. High quality predictions are done for ¹³C, and also ¹H predictions might be useful.²

Predictions by nmrshiftdb2 for nuclei other than ¹H are based on HOSE codes. These basically describe the neighbourhood of an atom in “spheres”. If two molecules have a similar neighbourhood, they will have the same HOSE code for a certain number of spheres. The more spheres used is also given. The more spheres can be used (maximum is 6) by the program, the better the prediction. Stereochemistry is supported in some cases, the result page will say “3D HOSE Code” if it was used, otherwise “2D HOSE Code” will be stated. For ¹H predictions apart from HOSE code, a system based on 3D descriptors and Support Vector Machines (“SVMs”) is offered. This can give better results than HOSE codes, but should be treated with care.²

• For using the “predict” function, you actually do not need to be logged on as a user. At the homepage of nmrshiftdb2 (http://www.nmrshiftdb.org) select the tab “predict”.

• You can either draw your structure or import various chemical file types using the icon. If you copy a SMILES to the clipboard you can directly paste it using Ctrl+V.

• By default, the prediction for ¹³C is preselected. If you would like to chose another nucleus, you can select it from the drop down list underneath the editor. By default, also only measured spectra are selected for the prediction, if you change to “calculated spectra”, also values of spectra which were only calculated are included. Submit the prediction by clicking on the “predict spectrum” icon.

• A typical result might look as the one given below:

¹ nmrhiftdb2 is developed by Stefan Kuhn and the NMR facility at University of Cologne. The precursor, NMRShiftDB, was the developed by the Steinbeck lab (EBI Cambridge). Further informations are available on the general homepage, http://www.nmrshiftdb.org/.
• If you have Java installed and want a dynamic display hit the symbol. Shifts can then either be highlighted in orange in the table by marking the corresponding atom in the structure editor, or vice versa. The type of shift values available/used for the prediction can be seen, if you move the cursor over the lower spectrum expansion. The interesting signal is highlighted in red once you touch it, and if you then hit the space key of your keyboard, an information window pops up. If a shift stems from only one compound, this is indicated by the number “1” in the last column (”values”).

• If you would like to change the molecule or the type of nucleus which is predicted, click on "modify molecule" button underneath the structure editor.

• The results can be printed in a more compressed layout by clicking the link ”print“ in the upper part of the prediction window.

• If the structure exists in the database, on the bottom of the prediction window, the link ”view structure in database“ allows you to see further information about the structure.
1.2 Search for structure

The search function is only explained for the most fundamental functions. For more detailed information (e.g. the so-called “expert mode” for searches) please check out the Help pages at the nmrshiftdb2 homepage.

- If you want to search for the NMR spectrum of a molecule, besides doing a prediction of the chemical shifts with the function mentioned in section 1.1 you can also directly do a search. Select the tab „search“ and scroll down to the section „search by structure“. Draw or import the structure as described in the previous paragraph.

- By default, the option „identity search“ is selected. The three possibilities of search types include "substructure search" (structures containing exactly the input structure, but also possible as a fragment of their complete structure), "similarity search" (structures with a similar structure to the molecule searched) or "identity search" (structures that are identical to the input structure). Click on the blue "search by structure" icon to submit. A typical result for an identity search might look like this:

- Besides the plotted spectrum, a table of chemical shifts and multiplicities (singulet = S, dublet = D, triplet = T and quadruplet = Q) is provided. If available, $J_{HH}$ couplings are listed in brackets. Note, that this identity search only yields results if the molecule is listed in the database.
• If you search for **similarity**, a list of similar structures (10 molecules per page) is presented who’s degree of similarity is compared via their Tanimoto coefficient (%). Likewise, doing a **substructure search** results in a list of all structures that contain the fragment defined as a substructure.

1.3 Search for chemical shifts or spectra

• Instead of structure, also **chemical shifts** can be the object of a search. This might be useful, if there is e.g. an unusual chemical shift in a spectrum. In such case, nmrshiftdb2 can provide structures that contain this chemical shift or similar ones and thereby help solving an unknown structure. In case, that you want to do a search for a specific chemical shift, mainly $^{13}$C shifts are advisable as the target. A search for single chemical shifts can be performed in the following manner:

The multiplicity of the signal (e.g. $S =$ singlet for quarternary carbons, $D =$ dublet for CH etc. for a $^{13}$C query) might be added with a semicolon as indicated in the right example. As default option, a „search“ spectrum search will be performed (for more information about subspectrum and complete spectrum search, please refer to the FAQ on the nmrshiftdb2 homepage). The result is listed as structures (10 per page) with similarity rating with respect to the chemical shift searched for.

• Another option is to do a **search for a complete or partial spectrum**, i.e. enter all or several shifts of the peaklist. Here, the same menu as described before for the chemical shift search is used. In this type of search, you can also upload the jcamp file of your spectrum (IUPAC format for spectra data) via the „upload file“ option underneath the field for the input of chemical shift values. If you decide to type or copy/paste the shift values of a peak list, please type/paste them into the field „input list“ in one of the following ways:
The chemical shifts need to be listed each in a separate line. Values can be entered with decimal comma or point. Optional, relative peak intensities can be entered, separated from the shift value by colon or semicolon. If the intensities are not between 0 and 1, they will automatically be scaled to this range.

1.4 How to use the semi-automated assignment function of nmrshiftdb2

A very useful function of nmrshiftdb2 is its ability, to automatically assign a spectrum to a given structure. Since this can be a valuable tool for the preparation of assignments for lab journals or reports, in the following, an explicit example will be discussed. The first two of the following paragraphs only apply if you work with the nmrshiftdb2 lab system (LIMS) installed at your institution. On the public instance, you choose the “Assignment” tab and proceed directly with the third step.

- As a first step, if using a local installation of nmrshiftdb2 with LIMS, for the assignment of spectra, you can revert to completed orders which you submitted to the LIMS. Therefore, you need to be logged on with your user name. You start out from a known order (“sample I.D.”): Select from your list of completed orders in the “NMR lab administration” window the corresponding, completed order. Open the order by clicking on its name. The order form (including structure and requested experiments, see right) will be visible now.

- Besides the download function, in the lower part of the displayed order form is a link for the assignment of the structure (“Assign spectrum”, see snapshot to the right). By clicking on this
link, you are directly transferred to the assignment module which allows to combine the already existent structure of your order with a peaklist extracted from the corresponding spectra. Note, that you should have prepared a peaklist (list of chemical shifts for all signals in your spectrum) with the aid of your processing software (SpinWorks or TopSpin, instructions for use see on the NMR lab homepage). Otherwise, you will have to enter the chemical shifts of the signals manually in the next step.

- If you started from an order processed by the lab system (LIMS), you can still change the structure of the molecule at this point by clicking on the "modify molecule" icon below the editor, if needed. **If you start out from the “Assignment” tab**, instead of the steps described before, you need to draw the structure first. The structure can be drawn either with the internal structure editor (Marvin JS) or it might be imported in various file formats (MDL Molfiles V2000, ChemAxon Marvin Documents, SMILES, ChemAxon Extended SMILES, SMARTS, ChemAxon Extended SMARTS, InChi, Name, CML, MDL Molfile V3000, MDL SDfile, ChemAxon Compressed Molfile, ChemAxon Compressed SDfile, XYZ). Also, you can copy-paste a structure as SMILES from ChemDraw or import one of your past structures via the “Import from structures history” button. Submit the structure by clicking on the “Submit molecule” icon.
Then, click on the button "Submit spectrum type" (you can choose between various nuclei, $^{13}$C is the default setting) and enter the chemical shifts of your compound. It is important to enter only one shift per line in the format "7.25" (use comma or a dot to separate decimal digits). If you want to add the relative signal intensity (e.g. as resulting from integration), as well, the entry for a line should look like "7.25; 0.3" or "7.25; 0.3". However, entering signal intensities is optional. If entered, it is important that all intensity values are in a range between 0 and 1.0. If you would like to enter multiplicities, e.g. for carbon signals, your entry should look like "77.0 T" or "124.0 S". Instead, you can simply read in the peak list in jcamp.dx format, which has been created in a program like e.g. SpinWorks: Click on the icon "Browse" and upload the peak list from your computer with "Import file". Finally, click on the button "Submit signals". This yields the following listing of pre-assigned chemical shifts:
• The next step is to assign the stored signals. By clicking on the icon “Do assignment” you can assign the chemical shifts entered before to atoms in the structure with the aid of the software. Once the window opens, nmrshiftdb2 suggests an assignment of the chemical shifts in the list for all atoms of the corresponding nucleus in the molecule. At this point, the assignment might be corrected based on experimental results. Also, the atom numbering can be changed (e.g. for a IUPAC-conform scheme). If you submit a $^1$H NMR dataset, you can add coupling constants at this point by clicking on the link “I want to enter coupling constants”. Otherwise, you confirm at this point the assignment by clicking on the icon “Submit assignments”.

• In the following screen, the assigned structure appears, and the assignment can be exported in various formats (e.g. html, pdf, rtf or in preformatted report styles of the corresponding research lab groups) by clicking on the link “Get this spectrum and its molecule as”. Thus, you can import, after doing the software-assisted assignment, your results into your lab report etc. At this point, you can also continue and submit your completed assignment to the (local) database, as is described in the following chapter.

### 1.5 Submitting assigned or non-assigned data to the (local) database, QuickCheck for assignment quality control

If you intend to submit a molecule to the local database, you might either start out from the assignment tab or a completed order, as is described in section 1.4. Or, you can directly select the “Submit” tab of the menu (you need to be logged on to your user account for this purpose).

• Initially, after starting the “Submit” menu, you need to draw (or import) the chemical structure of your molecule. The complete procedure for a submit includes five mandatory steps, two more steps (addition of references and attachment of spectra) are optional.
• The structure can be drawn either with the internal structure editor (Marvin JS) or it might be imported in various file formats (MDL Molfiles V2000, ChemAxon Marvin Documents, SMILES, ChemAxon Extended SMILES, SMARTS, ChemAxon Extended SMARTS, InChI, Name, CML, MDL Molfile V3000, MDL SDfile, ChemAxon Compressed Molfile, ChemAxon Compressed SDfile, XYZ). Also, you can copy-paste a structure as SMILES from ChemDraw or import one of your past structures via the “Import from structures history” button.

• Then, click on the button "Submit spectrum type" (you can choose between various nuclei, $^{13}$C is the default setting) and enter the chemical shifts of your compound. It is important to enter only one shift per line in the format "7.25" (use comma or a dot to separate decimal digits). If you want to add the relative signal intensity (e.g. as resulting from integration), as well, the entry for a line should look like "7.25; 0.3" or "7.25; 0.3". However, entering signal intensities is optional. If entered, it is important that all intensity values are in a range between 0 and 1.0. If you would like to enter multiplicities, e.g. for carbon signals, your entry
should look like “77.0  T” or “124,0  S”. Instead, you can simply read in the peak list in jcamp.dx format, which has been created in a program like e.g. SpinWorks: Click on the icon “Durchsuchen” and upload the peak list from your computer with “Import file”. Finally, click on the button “Submit signals”. This yields the following listing of pre-assigned chemical shifts:

• The next step is to assign the stored signals. By clicking on the icon “Do assignment” you can assign the chemical shifts entered before to atoms in the structure with the aid of the software. Once the window opens, nmrshiftdb2 suggests an assignment of the chemical shifts in the list for all atoms of the corresponding nucleus in the molecule. At this point, the assignment might be corrected based on experimental results. Also, the atom numbering can be changed (e.g. for a IUPAC-conform scheme). If you submit a $^1$H NMR dataset, you can add coupling constants at this point by clicking on the link. Otherwise, you confirm at this point the assignment by clicking on the icon “Submit assignments”.

• In the following screen (see next page), the assigned structure appears, and the assignment can be exported in various formats (e.g. html, pdf, rtf or in preformatted report styles of the corresponding research lab groups) by clicking on the link “Get this spectrum and its molecule as”. Thus, you can import, after doing the software-assisted assignment, your results into your lab report etc.
• In addition to the assignment, some more information is required to submit the data to the database. By clicking the icon “Add miscellaneous data”, for the field “Enter categories...” (drop-down list to the right) the entry “chemie koeln” is mandatory (!!!), since it describes the data as originating from the in-house database. Additionally, there are also mandatory fields for entering the temperature, solvent, field strength (resonance frequency for $^1$H) and type of assignment. Also, more information like e.g. name of the compound, CAS number, weblinks with information about the compound or keywords (e.g. ketone, ester, etc. – here you can also mark one or several keywords from a huge list) can be entered. Once you are done, click on “Submit data”.

• In case of non-published structures, the following step can be skipped. Here, literature references can be entered, including electronic links like doi.
• Also optional, but advisable, is the submission of spectral data together with the assignment. By clicking in “Attach jcamp-dx file”, your actual spectrum can be uploaded and attached to the dataset. Instead, you might also attach pdf files of spectra plots. This is not mandatory, but you should consider to add the data of a compound always, if you plan to submit the dataset to the inhouse database. If you only use the submit function as assignment aid, do not bother with it. At this point, your assignment is ready for submission to the database. If you would like to have your contribution ready for submission but do not want to transfer it yet (e.g. in case of a publication), you can still receive a rating of your assignment and submit the structure at a later point. In such case, you need to check the box “I want to keep this submission private” and mark the option “I want to use the CSEARCH robot referee” (only active for 13C data) to activate the QuickCheck. If you would like to submit the data to the (local) database, and, thereby, the referee, do not mark the “private” box.

Once you have completed all fields of interest, the assignment can be transmitted by clicking in “Write to database”. The consecutive window allows you to check and do final corrections before you submit the data to the database. After submission, you will be informed after ca. 20 min. that your assigned spectra are now queued for review (referee process). Here, the assignment quality control (QuickCheck) will be performed (if option was activated during submission) by aid of the Robot Referee of CSEARCH\(^3\) (13C-only) and the nmrshiftdb2 algorithm (HOSE code based, all nuclei). At this stage already, you can get the output for your report, by clicking on the tab “Personal Page” in the menu line. Now, you will see all non-reviewed spectra/structures on the left side.

\(^3\) CSEARCH and Robot Referee are developed by Prof. Wolfgang Robien, University of Vienna. For further information, see [http://homepage.univie.ac.at/Wolfgang_Robien/csearch_main.html](http://homepage.univie.ac.at/Wolfgang_Robien/csearch_main.html).
• By clicking on the molecule, whose assignment you would like to export (red arrow in snapshot), the complete assignment details (line spectrum, assigned peaklist and prediction) and the structure will be visible. To export the assignment, mark the “Download” tab to the right. Now, you can chose to export the assigned spectrum in the desired format selecting the option “Get this spectrum and its molecule as” and marking the desired export format in the pull-down list (e.g. “experimental section berkessel...”) and hitting the “request” icon.

• Once the rating for the QuickCheck has arrived (after a maximum wait of 2 hours), you can also see the evaluation of your assignment: Besides the assignment export option for a molecule, which was described in the previous section, the tab labelled “Additional Data” (see red label in snapshot) presents the results of the evaluation.
The rating (see snapshot below) might be used to revise the assignment, if necessary: CSEARCH presents the results of the Robot Referee with a colour code and a recommendation (green – accept, yellow – minor revision, red – rejected), nmrshiftdb2’s rating evaluates with points from 1 (poor) to 10 (accept).

• **Note**: If you checked the “private” mode for submission of your data, you need to activate the link “Show all my contributions” on your personal page, to select the desired molecule and then see the QuickCheck rating.