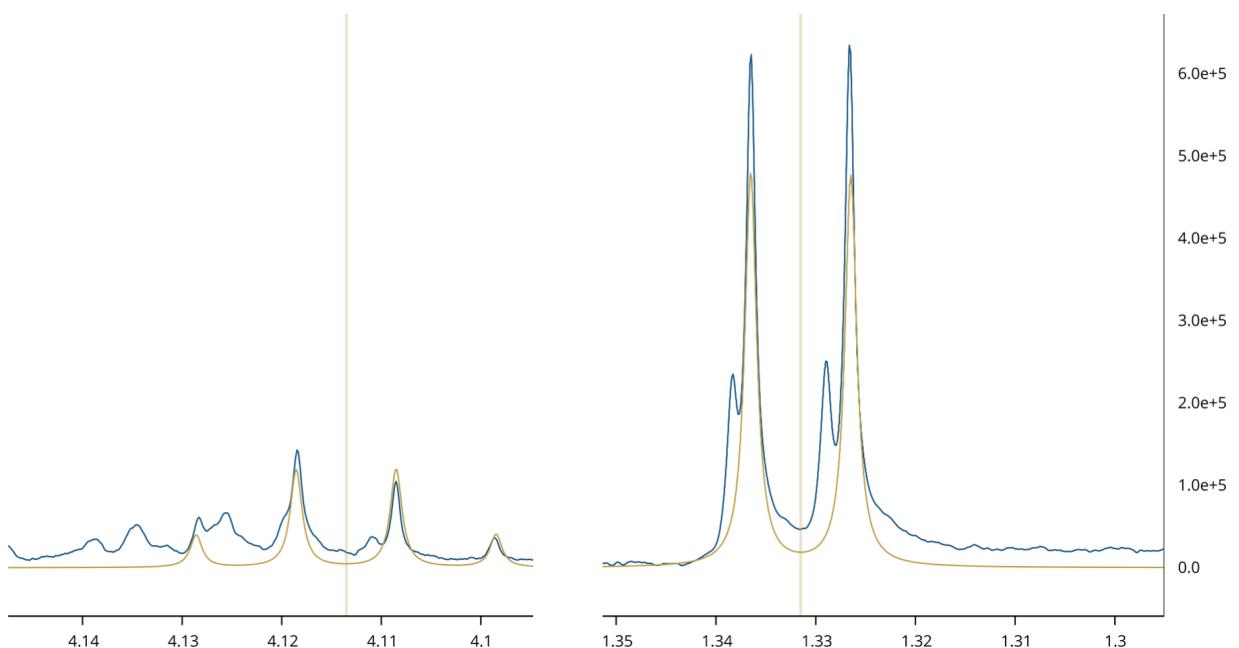


## AnalysisMetabolomics Profile by Reference Tutorial



# Introduction

This tutorial is designed to introduce you to CcpNmr AnalysisMetabolomics 3.2. In particular, the tutorial centres on profiling of 1D metabolomics spectra using reference data from a database.

It is assumed that you have some basic familiarity with the CcpNmr Analysis program (e.g. from having completed our Beginners Tutorial) and the theory behind the profiling of metabolomics spectra.

We recommend that you use the project provided in the **MetabolomicsTutorial** data directory provided by CCPN while working through the examples and problems. We are grateful to Dr Marie Phelan for providing the data. It is a subset of the data deposited in the [MetaboLights Database, entry MTBLS541](#).

Please note that the images shown are only representative and you may encounter minor differences in your setup.

## Contents

1. Loading data
2. Profile by Reference

## Start CcpNmr Analysis V3

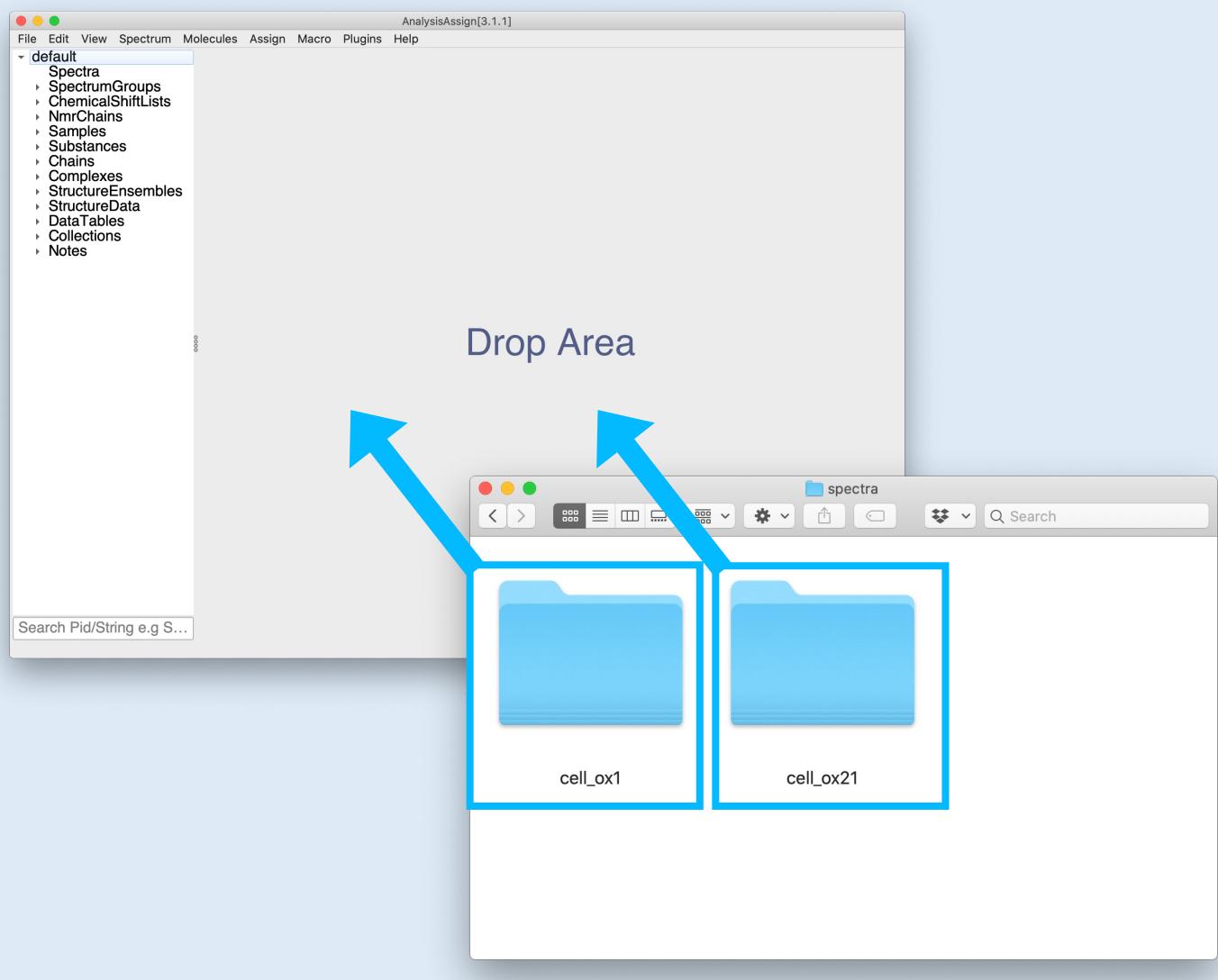
Apple users by double-clicking the *CcpNmrAnalysis* icon



Linux users by using the terminal command:  
*bin/assign*

Windows users by double-clicking on the *assign.bat* file

# Loading data



## 1A Load spectra

- Find the **spectra** folder in the Metabolomics Tutorial data directory.
- Select and drag the following three folders into the Sidebar or Drop Area:

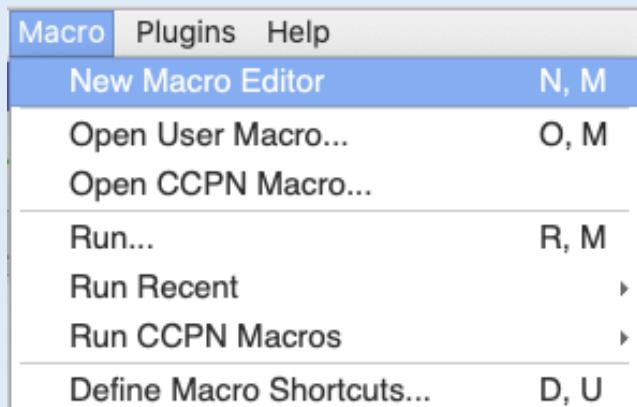
**cell\_ox21** (cell extract from cells grown at normal 21% Oxygen levels)

**cell\_ox1** (cell extract from cells grown at hypoxic 1% Oxygen levels)

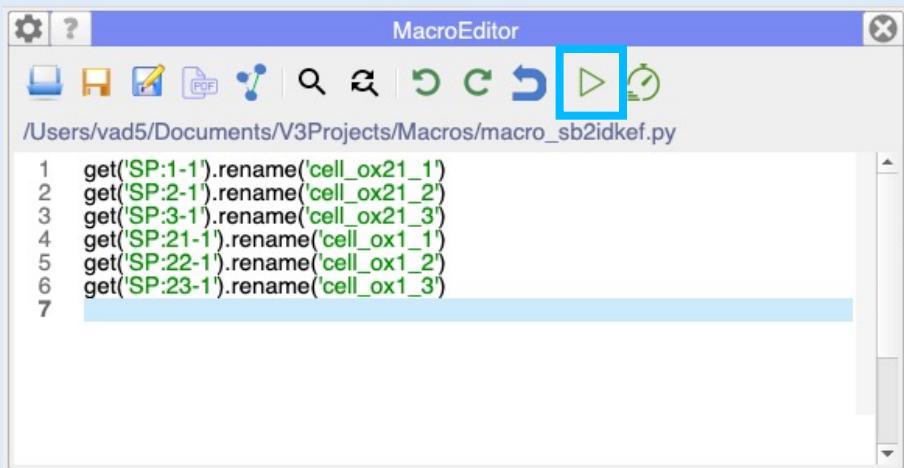
If you drop the folders into the Drop Area the spectra will open directly. If you drop the folders on the Sidebar the spectra will simply be visible in the **Spectra** section of the sidebar and can be opened in a SpectrumDisplay module at a later stage.



# Loading data



or Shortcut NM



## 1B Rename spectra

- Go to Main Menu → Macro → New Macro Editor

OR

- Use shortcut NM

This will open the MacroEditor module.

- Type or Copy/Paste in the following commands:

```

get('SP:1-1').rename('cell_ox21_1')
get('SP:2-1').rename('cell_ox21_2')
get('SP:3-1').rename('cell_ox21_3')
get('SP:21-1').rename('cell_ox1_1')
get('SP:22-1').rename('cell_ox1_2')
get('SP:23-1').rename('cell_ox1_3')

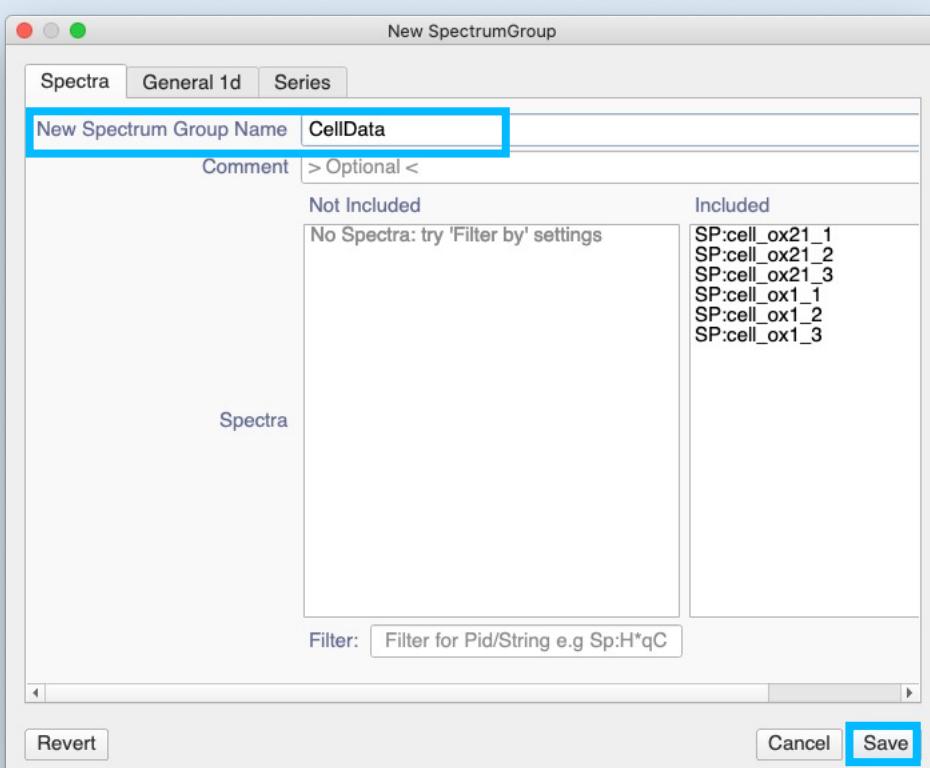
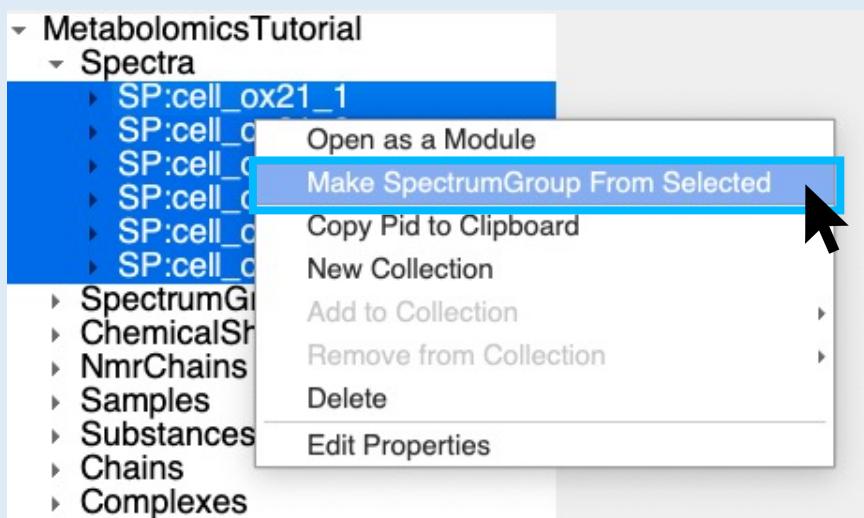
```

- Run the macro by pressing the green Play button
- Close the Macro Editor and Python Console which will have opened.

Your spectra should now be renamed in the Sidebar:

- ▼ Spectra
  - ▶ SP:cell\_ox21\_1
  - ▶ SP:cell\_ox21\_2
  - ▶ SP:cell\_ox21\_3
  - ▶ SP:cell\_ox1\_1
  - ▶ SP:cell\_ox1\_2
  - ▶ SP:cell\_ox1\_3

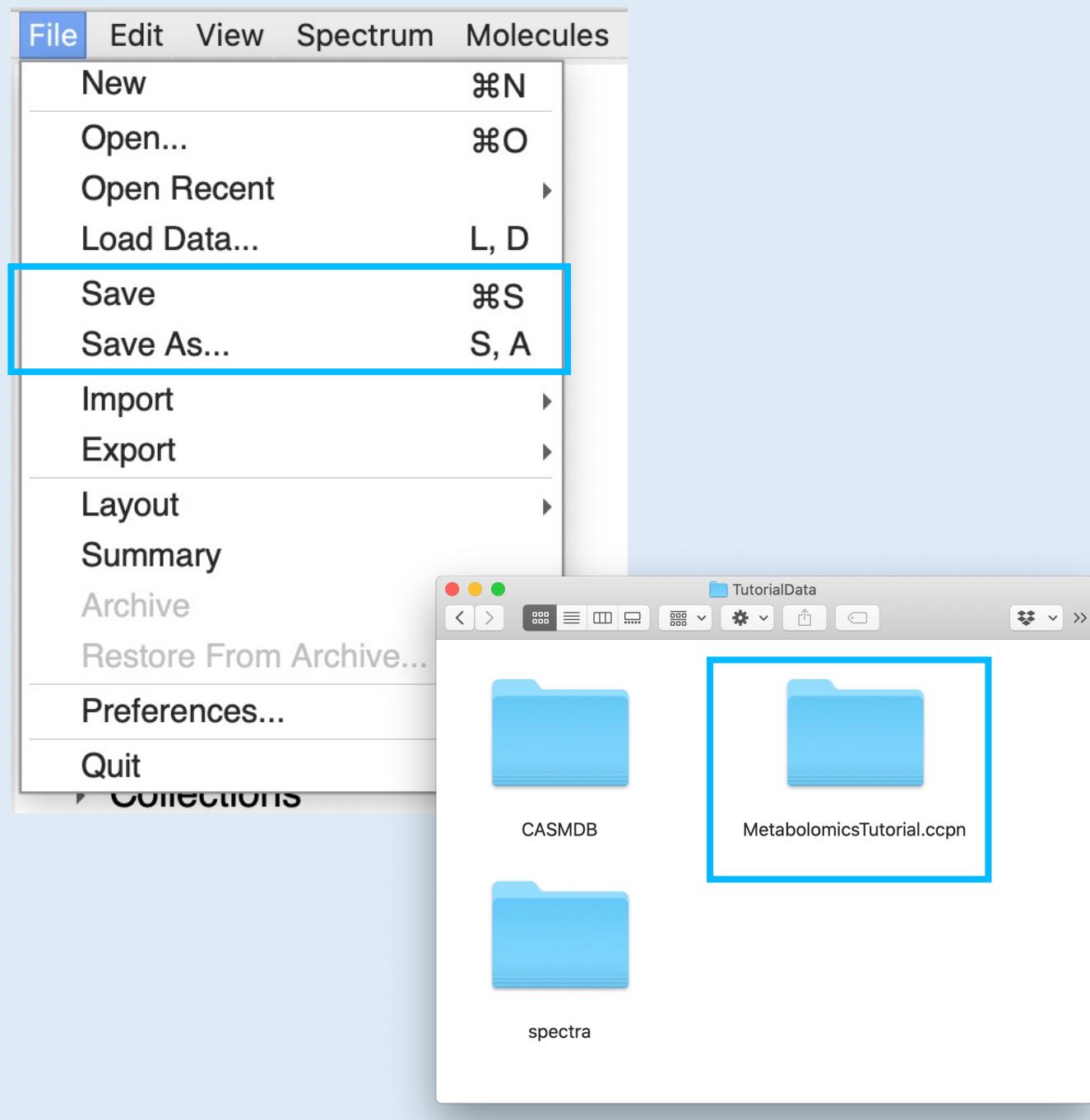
# Loading data



## 1c Create SpectrumGroup

- Select all your Spectra in the Sidebar (hold down the Shift key and select the top and bottom Spectrum)
- right-click and select **Make SpectrumGroup From Selected**
- In the pop-up you can give your SpectrumGroup a new name, e.g. **CellData**, if you wish. Then click on **Save**.

# Profile by reference



## 2A Save Project

The **Profile by Reference** feature can only be used if your initial temporary project has been saved, so there is a project directory in which information can be placed.

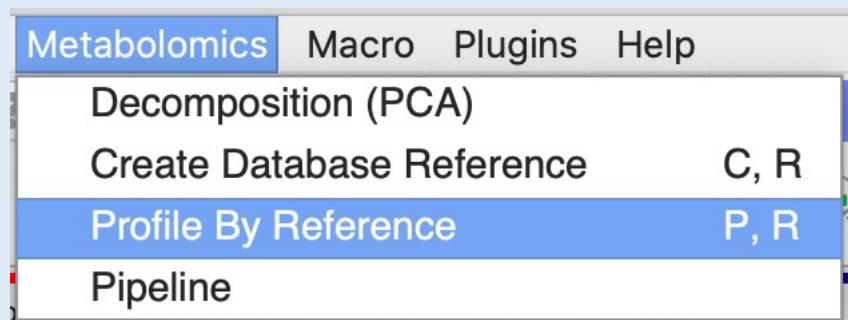
- If you haven't already saved your project, go to **Main Menu → File → Save** or **Save As...**

**OR**

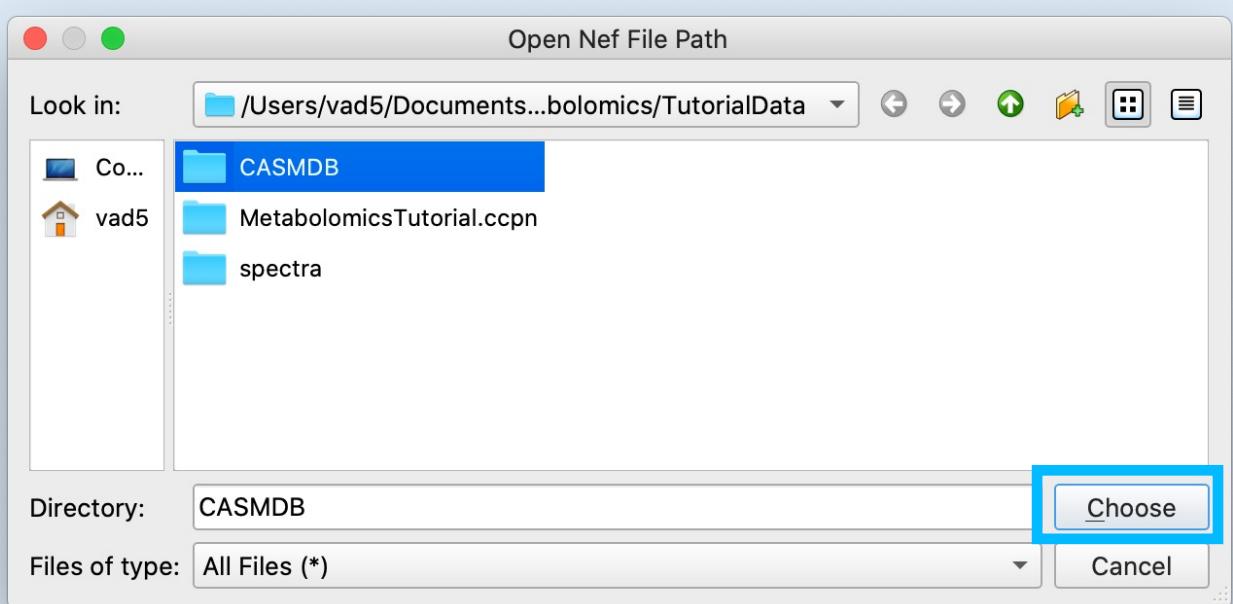
- Use shortcut **Ctrl/Cmd + S** or **SA**

Alternatively, load the **MetabolomicsTutorial ccpn** project provided.

# Profile by reference



or Shortcut  
PR



## 2B Load CASM-DB

- Go to Main Menu → Metabolomics → Profile by Reference

OR

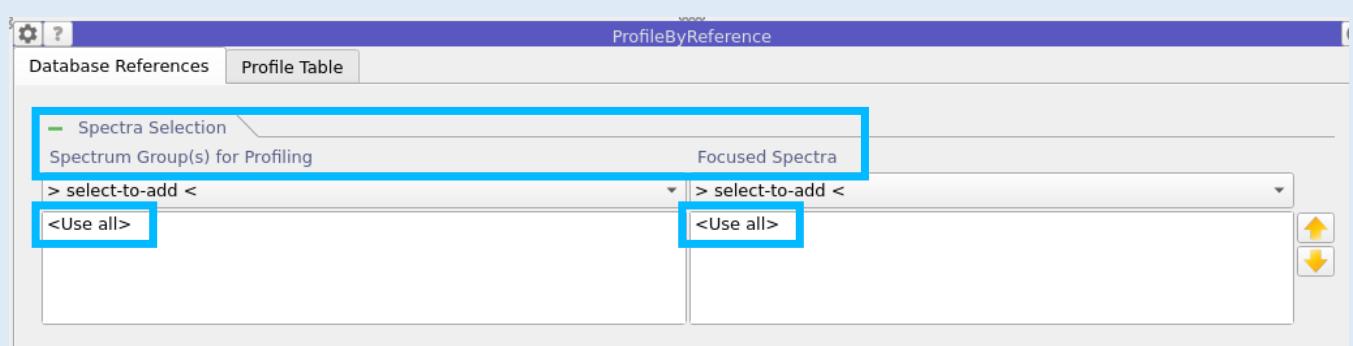
- Use shortcut PR

You may encounter a warning that this feature is still new and being tested.

If this is the first time you have used **Profile by Reference**, you will then be asked to select your database directory location.



- Select the **CASMDB** directory in the Metabolomics Tutorial data folder.



## 2C Spectra Selection

If the spectra are not already displayed in a SpectrumDisplay module:

- Open all the spectra in a Spectrum Display by **dragging** the spectra in the display area.

OR

- **Right-click → Open as Module.**
- In the **Profile by Reference** module, under **Spectra Selection**, make sure **<Use all>** is selected for both **Spectrum Group(s) for Profiling** and **Focused Spectra**.

**Spectrum Group(s) for Profiling** refers to the spectra which are included in the profile.

**Focused Spectra** selects which spectra are shown in the SpectrumDisplay module and for which spectra the current reference data will be saved in the **Profile Table**. This is covered in more detail in **part 2G**.

Initially, we want to save our reference to all spectra before fine tuning it for each spectrum individually.

# Profile by reference

The screenshot shows the 'Simulation Selection' interface. At the top, there's a 'Substance' section with a list of substances (e.g., (+)-abscisic\_acid, (+)-epicatechin, (+)-limonene, etc.). Below this is a 'Filter by Substance' dialog with the search term 'l-lactic\_acid' entered. A blue arrow points from the 'Search' button in this dialog to the corresponding row in the substance list below.

	name
1	(+)-lactic_acid

## 2D Simulation Selection

- Click **Filter by Substance** to open the substance filter menu.
- Type **l-lactic\_acid** in the search bar and click **search**.
- Click on **l-lactic\_acid** to open the simulation popup. In this popup there will be several options for simulations to reference from.
- Scroll down to **row #6 (spin\_system, GISSMO, 600.130)** and click on the row to generate the simulation and open the **Active Simulation Parameters**.

The screenshot shows a 'Simulation' table with the following data:

simulation_type	simulation_origin	spectrometer_frequency	metabolite_cc
peak_list	HMDB	499.620	None
peak_list	HMDB	500.000	None
peak_list	BMRB	-1.000	100
peak_list	BMRB	500.000	100
peak_list	BMRB	600.000	100
spin_system	GISSMO	600.130	None
spin_system	GISSMO	499.840	None

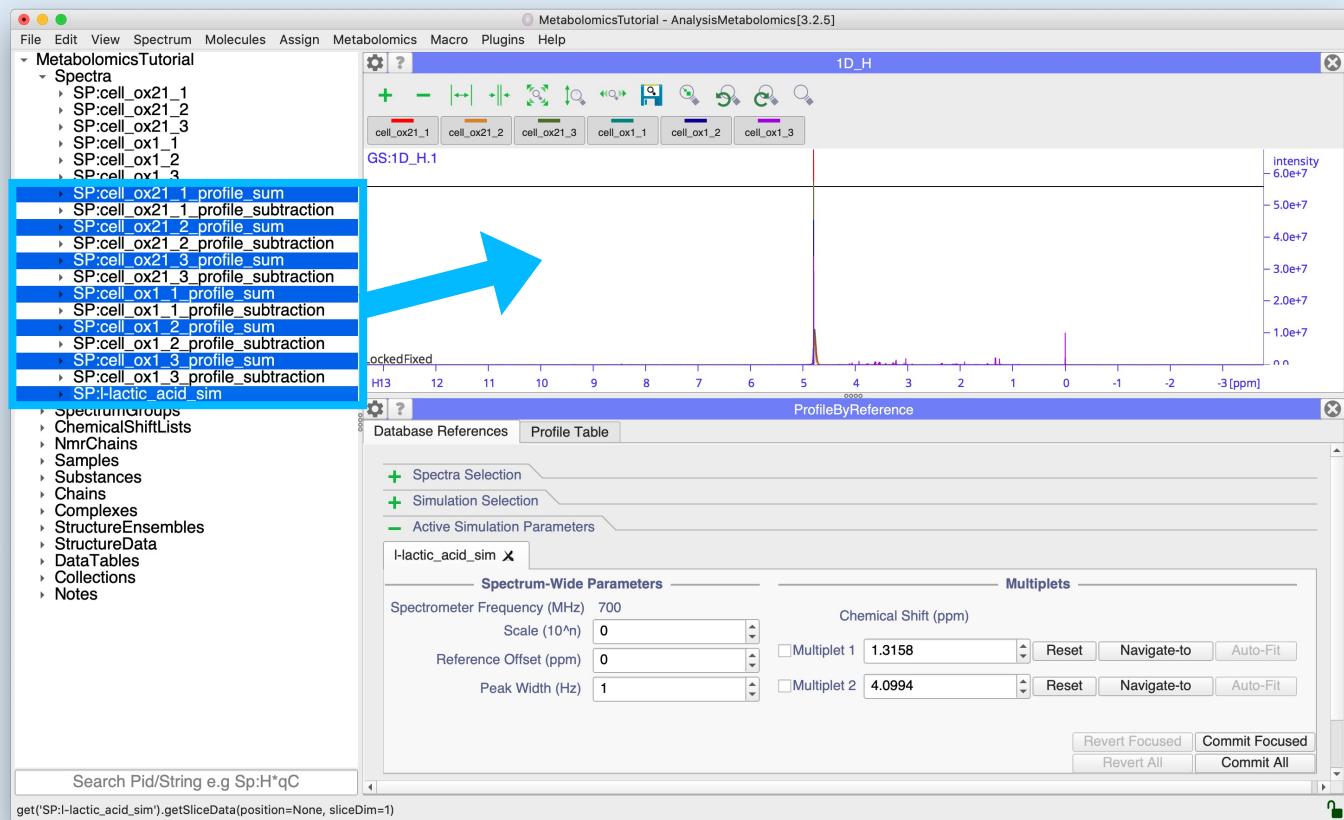
This may take a few seconds to load as the module generates the sum and subtraction spectra.

**Note:** The spectra used here were collected at 700MHz so only database references that can be simulated at 700MHz are suitable for comparison.

Simulations of the type **peak\_list** are fast to generate but cannot be extrapolated to other spectrometer frequencies.

Simulations of the type **spin-system** can be slower to generate but are field-independent and can account for strong coupling.

# Profile by reference



## 2E Substance and Profile Sum Spectra

In the sidebar there will now be 12 new spectra labelled **profile\_sum** and **profile\_subtraction** as well as one new spectrum **SP:I-lactic\_acid\_sim**.

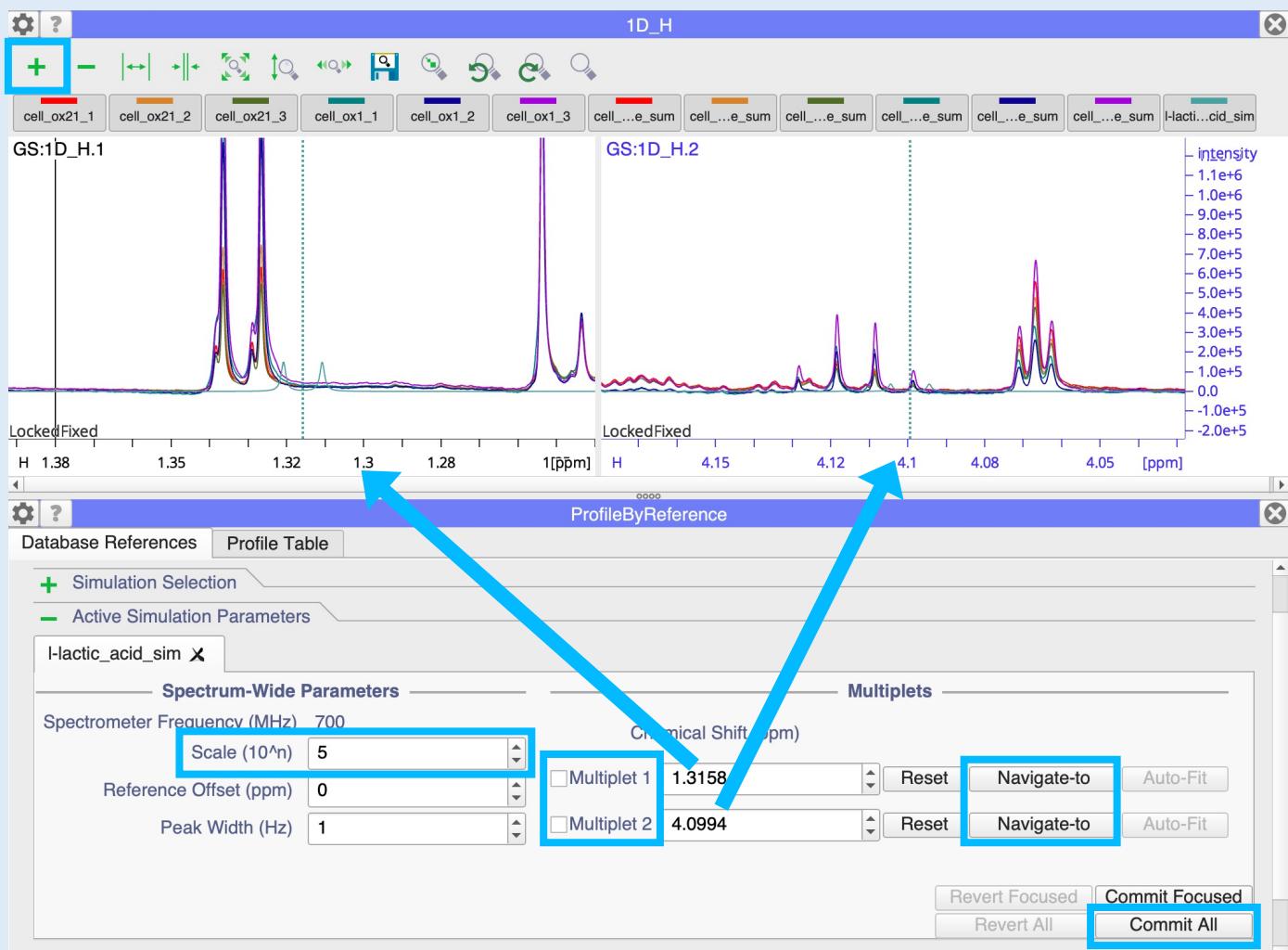
These are simulated spectra that have been generated by the module and are saved as .ndf5 files in the project **data/spectra** folder.

The **profile\_sum** spectra are the summed spectrum intensities of all the substance spectra produced by the module. The **profile\_subtraction** spectra are the intensities of the **profile\_sum** spectra minus the sample spectra.

- Drag all the **profile\_sum** spectra and the **SP:I-lactic\_acid\_sim** spectrum into the SpectrumDisplay.

These spectra will not contain any signal at the moment.

# Profile by reference



## 2F Adjust Active Simulation Parameters

- In the **Profile by Reference** module, scroll down to the **Active Simulation Parameters** frame and change the **Scale (10<sup>n</sup>)** value to 5.
- Click on the **Multiplet 1 Navigate-to** button to move the current strip's 1H axis to the chemical shift of Multiplet 1.  
You may need to zoom in to see the multiplet.
- Create a new strip with the **+** button at the top of the SpectrumDisplay.
- Click on the **Navigate-to** button belonging to **Multiplet 2**.
- Using the vertical line in the spectrum display or the **Chemical Shift (ppm)** spin box in the **Active Simulation Parameters** frame, align the multiplets to match the sample spectra.  
Approximately 1.33 for **Multiplet 1** and 4.11 for **Multiplet 2**.
- Click on **Commit All** to save these values to all samples, so you have a good starting point before fine tuning each sample.  
(Since all spectra are currently focused, you could also have clicked **Commit Focused**.)

# Profile by reference

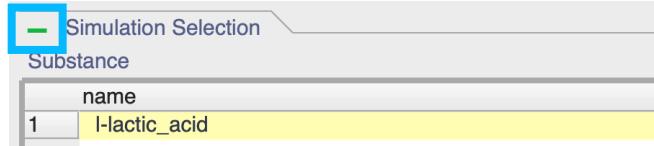


## 2G Adjust Active Simulation Parameters

- Click the yellow arrows next to the **Focused Spectra** list to change the focused spectra and toggle the spectra in the SpectrumDisplay on/off.
- Cycle through the focused spectra, and for each one, align the multiplet **Chemical Shift** and adjust the **Scale** using the vertical lines in the display or the spin-boxes in the **Profile by Reference** module.
- Click on **Commit Focused** to save the values for that sample/spectrum before moving on to the next one.

### Notes:

- You can also select specific combinations of focused spectra using the drop-down menu.
- The lactic acid doublet at 1.3ppm overlaps with signals from threonine. The intensities are therefore likely to be unreliable and it is best to set your **Scale** based on the quartet at 4.1ppm.
- You can minimise the **Simulation Selection** section of the module to have the **Spectra Selection** and the **Active Simulation Parameters** sections visible at the same time.



# Profile by reference

ProfileByReference

Database References Profile Table

Simplified Advanced

	spectrum	intensity
1	SP:cell_ox21_1	5.502
2	SP:cell_ox21_2	5.502
3	SP:cell_ox21_3	5.372
4	SP:cell_ox1_1	5.782
5	SP:cell_ox1_2	5.722
6	SP:cell_ox1_3	5.972

.CSV  
.tsv  
.json

Load Profile Export Profile

ProfileByReference

Database References Profile Table

Simplified Advanced

	simulated_spectrum	origin	spectrum	substance	intensity
1	nef_nmr_spectrum_L-lactic_acid_5`1`	database	SP:cell_ox21_1	SU:L-lactic_acid.metabolite	5.502
2	nef_nmr_spectrum_L-lactic_acid_5`1`	database	SP:cell_ox21_2	SU:L-lactic_acid.metabolite	5.502
3	nef_nmr_spectrum_L-lactic_acid_5`1`	database	SP:cell_ox21_3	SU:L-lactic_acid.metabolite	5.372
4	nef_nmr_spectrum_L-lactic_acid_5`1`	database	SP:cell_ox1_1	SU:L-lactic_acid.metabolite	5.782
5	nef_nmr_spectrum_L-lactic_acid_5`1`	database	SP:cell_ox1_2	SU:L-lactic_acid.metabolite	5.722
6	nef_nmr_spectrum_L-lactic_acid_5`1`	database	SP:cell_ox1_3	SU:L-lactic_acid.metabolite	5.972

## <sup>2</sup>H Profile Tables

- Click on the **Profile Table** tab to see the committed measurements for the profile.

By default, the **Simplified** tab will be visible. This table contains a record of the profiling done so far and also highlights in orange where any uncommitted changes have been made.

- Click the **Advanced** tab to see a more detailed record of the profile, which includes the simulation parameters.
- Click the **Export Profile** button to save the simplified profile table to another location (as an Excel File, .csv, .tsv or .json).

Either table can also be exported with **right-click → Export Visible Table**.

	spectrum	intensity
1	SP:cell_ox21_1	5.502
2	SP:cell_ox21_2	5.502
3	SP:cell_ox21_3	5.372
4	SP:cell_ox1_1	5.782
5	SP:cell_ox1_2	5.722
6	SP:cell_ox1_3	5.972

Filter...  
Copy clicked cell value  
Delete Selection  
Clear Selection  
Export Visible Table  
Export All Columns  
Remove Clicked Row  
Remove Clicked Column  
Remove Clicked Cell Value

# Profile by reference

The screenshot shows the software's main window with the 'Database References' tab selected. In the 'Profile Table' section, there is one entry: spectrum SP:cell\_ox21\_1 at 5.502 ppm. On the right, the 'Active Simulation Parameters' section shows 'I-lactic\_acid\_sim' is active. Below it, 'Spectrum-Wide Parameters' include 'Spectrometer Frequency (MHz)' set to 700. The bottom left shows the 'Simulation Selection' section with a search bar containing 'threonine'. The 'Simulations in Profile' table lists one simulation: spectrum\_pid 1, simulation\_type SP:I-lactic\_acid\_sim, spin\_system GISSMO.

spectrum_pid	simulation_type	spin_system	simulation_origin
1	SP:I-lactic_acid_sim	spin_system	GISSMO

## 2I Add a second Substance

- Return to the **Database References** tab.
- Close the **I-lactic\_acid\_sim** tab in the **Active Simulation Parameters** section, by clicking the cross next to the tab title.
- Expand the **Simulation Selection** section if you had previously minimised it.
- Reset the search parameters by clicking the **Reset** button. (You may need to reopen the search menu by clicking the **Filter by Substance** button.)
- This time search for **I-threonine**.

Note that you can also use alternative search parameters as shown above if you don't know the full name of your compound.

- Select another spin-system simulation to profile with.

The screenshot shows the 'Simulations in Profile' table. It lists several simulations, with the last two rows highlighted in blue. The columns are: simulation\_type, simulation\_origin, spectrometer\_frequency, and metabolite\_cc. The highlighted row (row 6) has simulation\_type 'spin\_system', simulation\_origin 'GISSMO', spectrometer\_frequency '499.840', and metabolite\_cc 'None'.

simulation_type	simulation_origin	spectrometer_frequency	metabolite_cc
peak_list	UNIVID	500.000	None
peak_list	BMRB	-1.000	100
peak_list	BMRB	600.000	100
peak_list	BMRB	500.000	0.500
peak_list	BMRB	500.000	100
spin_system	GISSMO	499.840	None
spin_system	GISSMO	600.130	None

- Drag the new **SP:I-threonine\_sim** spectrum into the display.
- Repeat steps 2F-I.

Gradually build up your profile in this way. Further compounds you may wish to look for in these spectra are I-alanine and I-phenylalanine.

## Contact Us

**Website:**

[www ccpn.ac.uk](http://www ccpn.ac.uk)

**Suggestions and comments:**

[support@ccpn.ac.uk](mailto:support@ccpn.ac.uk)

**Issues and bug reports:**

<https://forum.ccpn.ac.uk/>

## Cite Us

Skinner, S. P. et al. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR* 66, (2016)