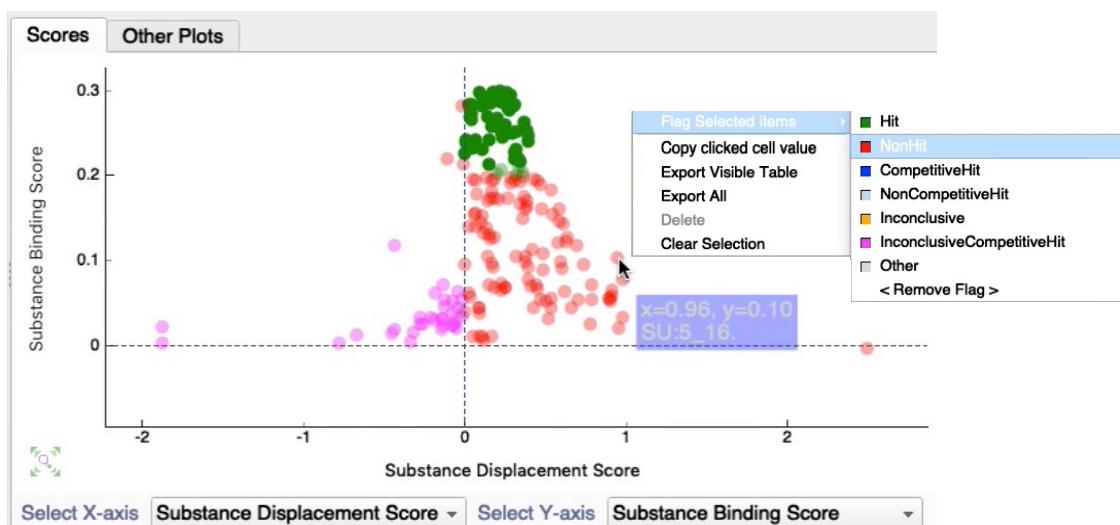


AnalysisScreen Hit Analysis Tutorial



Introduction

This tutorial will show the usage of the Hit-Analysis module in CcpNmr AnalysisScreen Version 3.1.

We strongly recommend that you read the Introduction to familiarise yourself with our terminology and workflows. Not all users may wish or need to work through all sections and using the projects provided it is easy to skip between different sections.

It is assumed that you have some basic familiarity with the program, e.g. from having completed our Beginners Tutorial.

You will need to use the data located in the **/data/ScreenTutorialMarch22** directory of the CcpNmr V3 examples data which you can download from the CCPN website

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

Contents

Introduction

Part I: Data Processing using Reference Singletons

- Import from Excel Files *(Section 1)*
- Manual Peak Matching from Reference Singletons *(Section 2)*
- Automated Peak Matching from Reference Singletons *(Section 3)*

Part II: Hit Analysis

- Binding Scores and Calculation Engines *(Section 4)*
- Plots and Filtering *(Section 5)*
- Exporting Data *(Section 6)*

Part III: Working with Reference Mixtures

- Create/update a NEF file *(Section 7)*
- Import from Excel and NEF Files *(Section 8)*
- Automated Peak Matching from Reference Mixtures *(Section 9)*

Part IV: Compare Screens

- Using data with multiple CPMG delays *(Section 10)*

Start CcpNmr Analysis V3

- Apple users by running Screen on the Launcher
- Unix users by using the terminal command: *bin/screen*
- Windows users by double-clicking on the *screen.bat* file

Disclaimer

Datasets used for this tutorial are randomly generated and don't have any biological significance. All spectra shown are synthetic and for demonstration purposes only. All compound names are randomly chosen and might have incorrect chemical properties or not be represented by the linked spectra.

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

Introduction

Getting started, basic operations

Sidebar

All data contained in a project, such as spectra and peak lists are located in the sidebar. **Double-clicking** on an item will open its properties popup.

Spectrum Display

A Spectrum Display can contain multiple overlaid spectra which share the same axes. To show/hide a single spectrum, click on its spectrum toolbar button. If you close a display, you can open a spectrum by **dragging and dropping** it into the drop area from the sidebar or by **right-clicking** on a sidebar item and selecting **Open as module**. You can also add additional spectra to a spectrum display module or drag several spectra into the drop area together to open them simultaneously.

Mouse

- Pan → **Left-drag** in display
- Zoom in/out → **Scroll wheel** in display
- Context menu → **Right-click**
- Select a peak → **Left-click** on a peak symbol “X”
- Move a peak → select first, then **middle-click and drag**

Two-Letter Shortcuts

Press the first letter on your keyboard e.g., **M**, followed by the second letter, e.g., **K** (case insensitive). Press **Esc** to cancel the first letter.

Common in this tutorial:

- | | |
|--------------------|---------------------------------------|
| SE | → Snap to Extremum the selected peaks |
| HA | → Open the Hit Analysis GUI Module |
| PI | → Open the Pipeline GUI Module |
| MC | → Clear all marks |
| Space-Space | → Open the Python console GUI Module |

For more commands and operations

Main Menu → Help → Tutorials → Beginners Tutorial

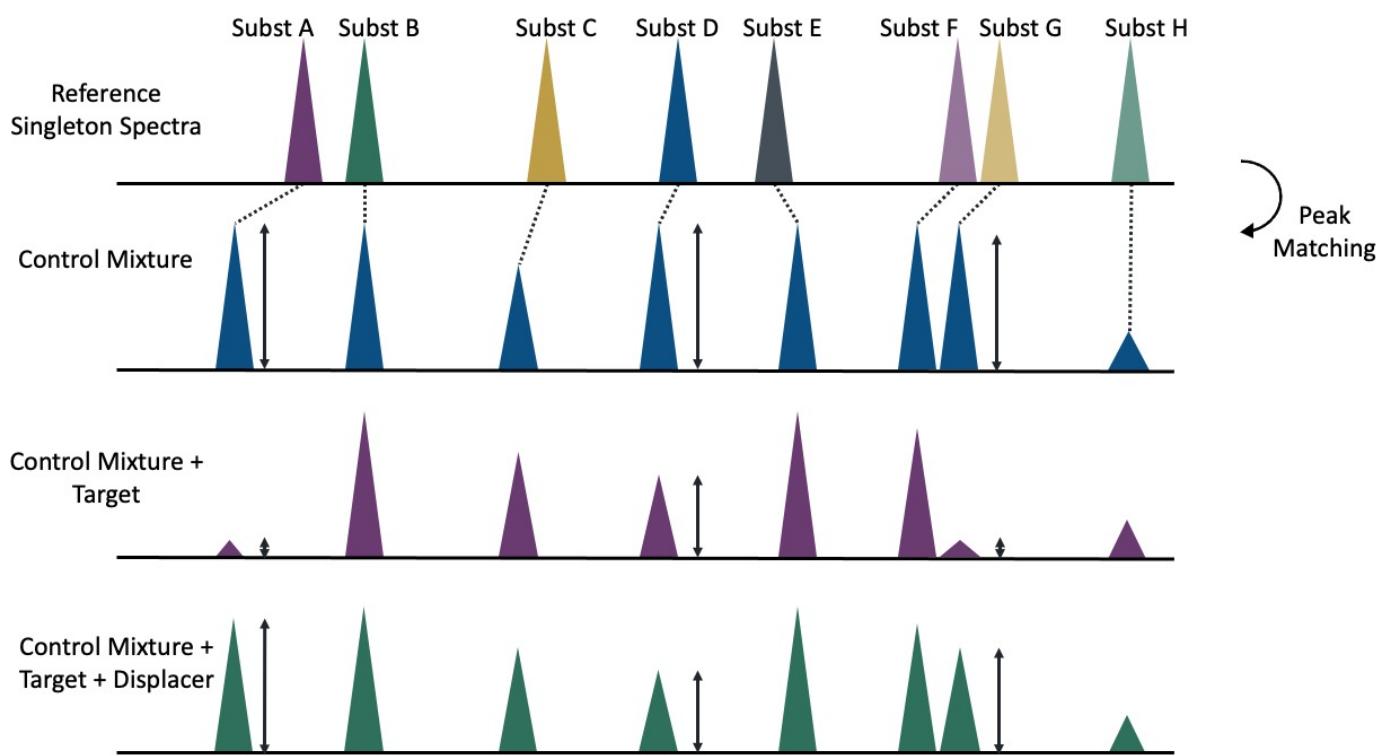
OR

Main Menu → Help → Show Shortcuts

Introduction

Schematic representation of the Screen analysis workflows

In this tutorial you will analyse several ^{19}F datasets but similar steps and tools in AnalysisScreen can also be applied to other experiment type analyses, such as ^1H relaxation-edited, WaterLOGSY and STD experiments.



We assume that you start with a set of individual spectra for each Substance (or compound) in your library. We refer to these as **Reference Singleton Spectra**.

Each screening experiment will start with the recording of a set of **Control Mixture Spectra**, a control spectrum of only the **Substance Mixtures**. Typically some of the peaks in the Control Mixtures will move relative to the Reference Singleton Spectra. The first task, therefore, is to match the peaks from the Reference Singleton Spectra to those in the Control spectra so that we know which peak in the Control Mixtures belongs to which Substance. This **Peak Matching** step can be done manually (Section 3) or automatically (Section 4). If done automatically, a **Peak Matching Score** will help you find matches that might need to be checked/corrected manually before you proceed.

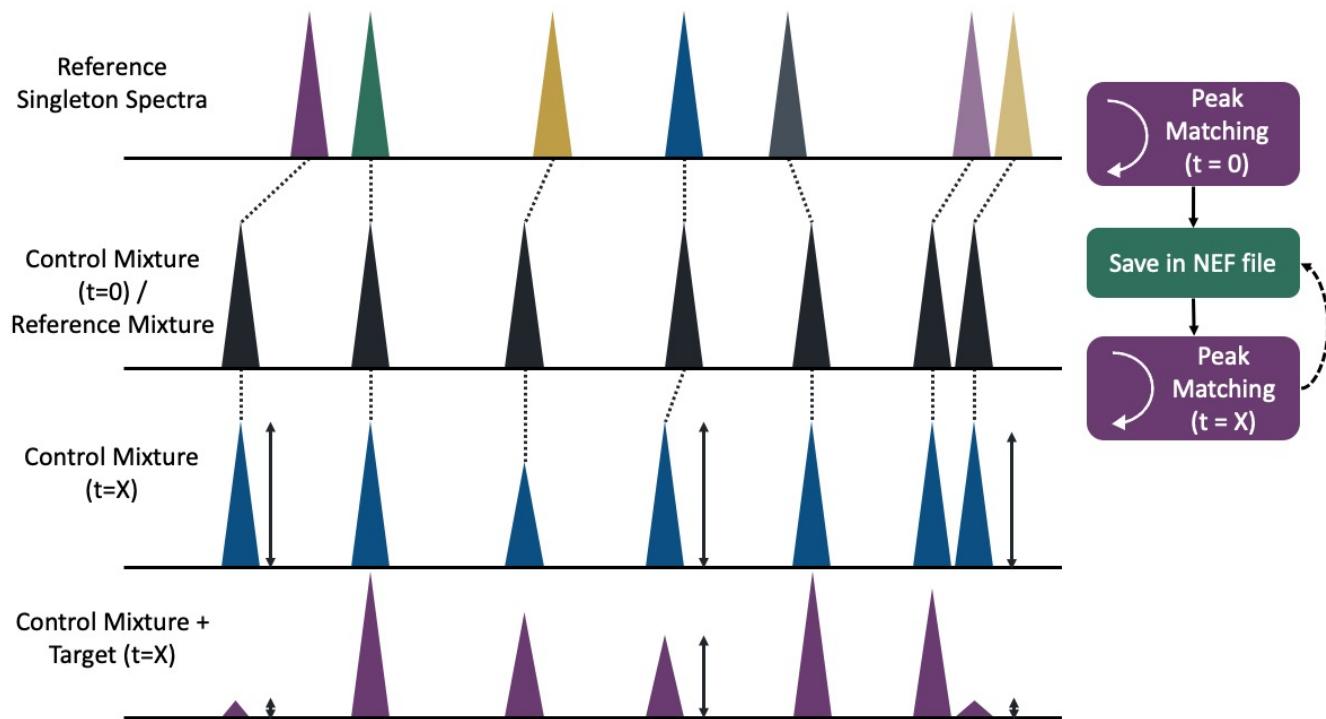
In addition to your Control Mixture Spectra you will have spectra where you have added a **Target** molecule and perhaps also a **Displacer** (also referred to as a **Competitor**). These spectra are in fact also included in the Peak Matching step, so that you end up creating a **Screening Dataset** in which the peaks are matched across all spectra and to a particular Substance.

Now you can use the **Hit Analysis module** to calculate the **Peak Binding Score** from the Control and Target spectra (exactly how is experiment type dependent, Section 6). Then inspect, classify and flag your Hits (Section 7) and export them (Section 8). Note that Binding Scores are also provided per Substance and per Sample.

Introduction

Recurring Screen analysis workflow

If you are repeating your screens with the same library of Substances and different Targets, then you can make the automated Peak Matching step faster, more accurate and more reliable by using previous Control Mixtures as **Reference Mixtures**.



Typically, there will be more differences between the Reference Singleton Spectra and the first Control Mixture spectrum you record, than between the Control Mixtures recorded at different time points.

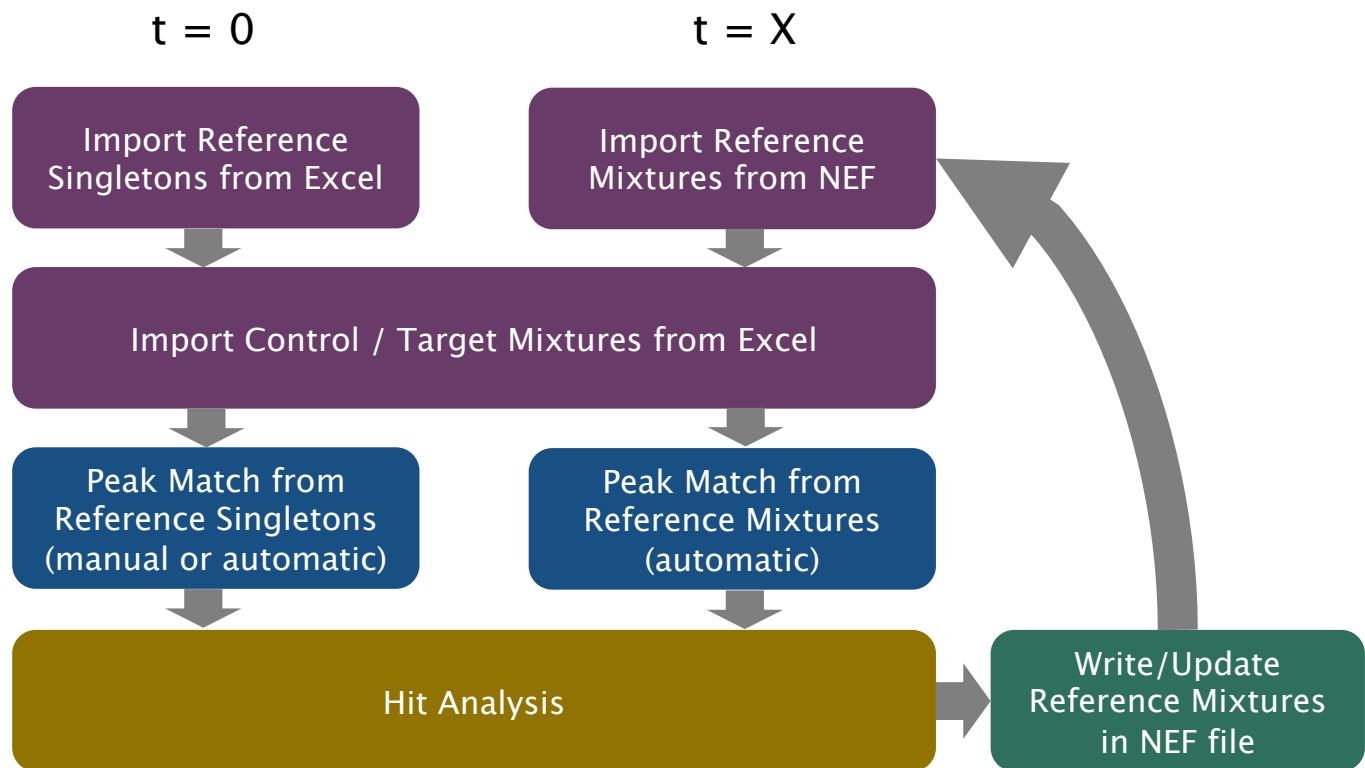
Therefore, we recommend that after your first Peak Matching (at time = 0), you save your Control Mixture data in an NMR Exchange Format (NEF) file and use this as a set of **Reference Mixtures** for your future screens. If your Control Mixtures continue to change over time, you can save each Control Mixture as the Reference Mixture for the next screen or you could create a new Reference Mixture NEF file once every 6 months or year, depending on how stable your Substance Mixtures are.

See the next page for an outline of the resulting workflow.

Introduction

... continued

The following diagram shows the steps conducted during the first analysis of your screening data at time = 0 and then of later analyses if you use previous Control spectra as Reference Mixtures, save in a NMR Exchange Format (NEF) file.



Introduction

CcpNmr AnalysisScreen Nomenclature

Sample

A CcpNmr object containing information about the physical NMR sample, e.g., pH, ionic strength etc.
CcpNmr links: Sample component, Spectrum (e.g.: the spectrum Control, Target...)

Sample component

A CcpNmr object containing information about the Substance in the physical NMR sample (e.g., concentration).

CcpNmr links: Substance

Substance

A CcpNmr object containing information about a molecule, (e.g., a small molecule and its general properties such as: SMILES, MW etc.).

CcpNmr links: Sample component, Spectrum (e.g.: the Singleton Spectrum)

SpectrumGroup

A CcpNmr object containing a collection of spectra.

CcpNmr links: Spectra

Control (spectrum)

The spectrum recorded at time X for a sample containing one or multiple substances prior the addition of a biological target.

CcpNmr links: Sample

Target (spectrum)

The spectrum recorded at time X for a sample containing one or multiple substances plus a biological target.

CcpNmr links: Sample

Displacer (spectrum)

The spectrum recorded at time X for a sample containing one or multiple substances plus a biological target and a known binder. Also referred to as a “displacer”.

CcpNmr links: Sample

Reference Mixture

The spectrum recorded for a sample containing multiple substances. Its peaks and their annotations are used as a template and to identify substances in future screening analyses.

CcpNmr links: Substances

Reference Singleton

The spectrum recorded for only one substance.

CcpNmr links: Substance

Binding Substance

The substance linked to a spectrum (reference) whose peaks have been matched to the spectral peaks (Control – Target) denoting a binding event.

Peak Match

The virtual linkage between a Reference – Control – Target (– Displacer) peak in a Screening Dataset.

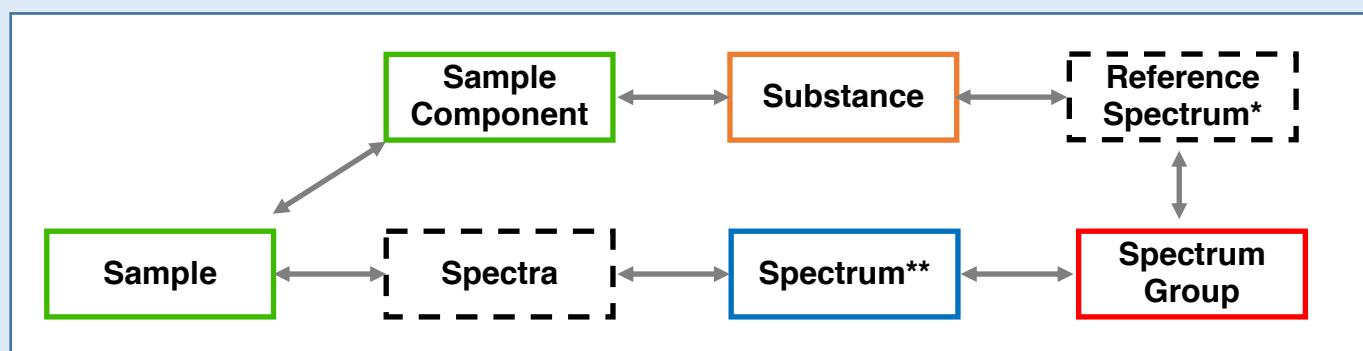


Figure showing CcpNmr AnalysisScreen object links

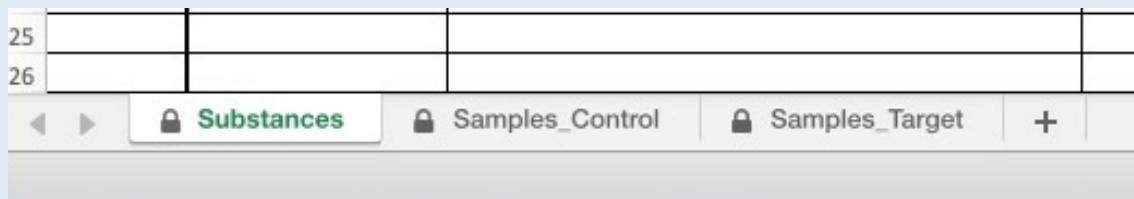
The schematic diagram shows how objects are linked in CcpNmr AnalysisScreen.

*Reference Singleton; **Control, Target, Displacer, Reference Mixture

Import Data from Excel

The program can read .xls or .xlxs files with multiple sheets that include the words **Sample** or **Substance** in the sheet name.

You can create files that contain either the Substance or Sample page or both.



1A Overview

Open a new Excel file or find a template in:

[ScreenTutorial/LookupTemplate.xls](#)

Mandatory Sheet names:

Title must start with “Substance” or “Sample”

Mandatory columns:

Substance sheet:

SubstanceName

Sample sheet:

SampleName

	A	B	C	D	E	F	G	H	I	
1		substanceName	spectrumPath	spectrumGroupName	experimentType	comment	smiles	synonyms	molecularMass	empirical
2	0	C101	C101	References	19F	OCCOC[C@@]1[C(=O)N]=N		()	160.151	
3	1	C102	C102	References	19F	COCCN1C[C@@H](CC1=O)[NH2]C[NH3]		()	189.255	
4	2	C103	C103	References	19F	N=[N]=N[C@]1[C#N]CCO[C@@H]1OC(=O)C		()	197.171	
5	3	C104	C104	References	19F	N=[N]=NC(=O)[C@@H](n1nnn(c1=S)C(=O)[C@@H]1C[C@@H]1Br)C		()	347.172	
6	4	C105	C105	References	19F	N=[N]=N[C@@H]([C@H](n1nnn(c1)c1ccccc1N1C(=O)C=CC1=O)C)C		()	338.344	

1B Create the Substance Sheet

The first sheet, **Substance**, can contain metadata associated with small molecules whose spectra, for example, have been used as references in a screen.

- Place the Lookup file template from the **ScreenTutorial** directory into the directory containing your spectra (**ScreenTutorial/19F/dataset1/data** if using the tutorial data).
- Open the template and fill in the **substanceName** column. This is the only mandatory column to fill in.
- Copy and paste the following:

C101
C102
C103
C104
C105

	A	B	C	D	E
1	substanceName	spectrumPath	spectrumGroupName	experimentType	
2	0	C101	C101	References	19F
3	1	C102	C102	References	19F
4	2	C103	C103	References	19F
5	3	C104	C104	References	19F
6	4	C105	C105	References	19F

1c Add reference spectrum information

To include the Substance reference spectra, you need to insert the **spectrumPath** (AnalysisScreen will recognise any spectrum format and you do not need to include any file extensions such as .ucsf, ndf5, .ft etc.).

You have three options:

1. If all the spectra files are located in the same directory as the lookup file, insert only the file names as above.
2. If the spectra are located in a subdirectory, insert the directory name first followed by a slash and the filename (the relative path starting from the Excel file), e.g. **references/C101**
3. If the spectra files are located in a completely different location, insert the full path, e.g. **/Users/username/Desktop/data3/MySpectra/C101**

For Bruker files, you can insert the path to the "r" file:

~ScreenTutorial/19F/dataset_1/data/pdata/1/1r

For clarity, we recommend keeping all the files in the same directory together with the Excel lookup file.

- Insert the **spectrumGroupName**; e.g. **References**. This will create a **Spectrum Group** with that name and place the spectra into it.
- Insert the **experimentType**. For these 1-dimensional 19F spectra, simply type **19F** into the cell.

Import Data from Excel

	A	B	C	D	E	F	G	H	I
1	substanceName	spectrumPath	spectrumGroupName	experimentType	comment	smiles	synonyms	molecularMass	em
2	0 C101	C101	References	19F		OCCOC[C@@]1(CO1)N=[N]=N	()	160.151	
3	1 C102	C102	References	19F		COCCN1C[C@@H](CC1=O)[NH2]C[NH3]	()	189.255	
4	2 C103	C103	References	19F		N=[N]=N[C@]1(C#N)CCO[C@@H]1OC(=O)C	()	197.171	
5	3 C104	C104	References	19F		N=[N]=NC(=O)[C@@H](n1nnn(c1=S)C(=O)[C@@H]1C[C@@H]1Br)C	()	347.172	
6	4 C105	C105	References	19F		N=[N]=N[C@@H]([C@H](n1nnn(c1)c1ccccc1N1C(=O)C=CC1=O)C)C	()	338.344	

1D Add Substance metadata

- The **comment** column will store any textual information about the substance. Avoid using letters with accents etc. (e.g. é, ä, œ, þ, ß, ç, ...) as these may cause problems when trying to save your CCPN project.
- If you enter the **smiles** for your substances, the program will automatically generate the structures inside the software. For the tutorial, copy and paste these SMILES:

OCCOC[C@@]1(CO1)N=[N]=N
 COCCN1C[C@@H](CC1=O)[NH2]C[NH3]
 N=[N]=N[C@]1(C#N)CCO[C@@H]1OC(=O)C
 N=[N]=NC(=O)[C@@H](n1nnn(c1=S)C(=O)[C@@H]1C[C@@H]1Br)C
 N=[N]=N[C@@H]([C@H](n1nnn(c1)c1ccccc1N1C(=O)C=CC1=O)C)C
- In the **synonyms** column you can insert the chemical name of the substance and again select **Match Destination** formatting
- All the following columns contain the substance chemical properties. Fill them if you want to display them within the software.
- Save the file.

A fully completed lookup file is provided at

[ScreenTutorial/19F/dataset1/data/lookup_19F_dataset1.xls](#).

	A	B	C	D	E	F	G	H	I	R
1	SampleName	spectrumGroupName	spectrumPath	spectrumName	experimentType	comment	pH	ionicStrength	sampleComponents	
2	0	Control_1	Control	Control_1	Control_1	19F	9.5	8	C101,C102,C103,C104,C105	
3	1	Target_1	Target	Target_1	Target_1	19F	9.5	8	C101,C102,C103,C104,C105	
4	2	Displacer_1	Displacer	Displacer_1	Displacer_1	19F	9	7.5	C101,C102,C103,C104,C105	
5										

1E Create the Samples Sheet

The next sheet in the template is **Samples**. This can contain metadata associated with particular samples, e.g. in a screening trial the sample could contain lots of spectra recorded with different experimental conditions. The only mandatory column is the **sampleName** column.

- Insert the **sampleName** in the first column, e.g. **Control_1**

The next three columns are specific to the spectra recorded for this sample:

- Insert the **spectrumGroupName**, e.g. **Control**, if you want the spectrum to be included in a Spectrum Group
- Insert the **spectrumPath**, e.g. (see the section **1C** for how to insert the spectrum path)
- Insert the spectrum **experimentType**, e.g. **19F**
- Fill in the **sampleComponents** column towards the end of the sheet. Insert the names of the components (**Substances**) that are present in the sample. In the case of a mixture containing components 1 to 5, insert them as a comma-separated list without spaces:

C101,C102,C103,C104,C105

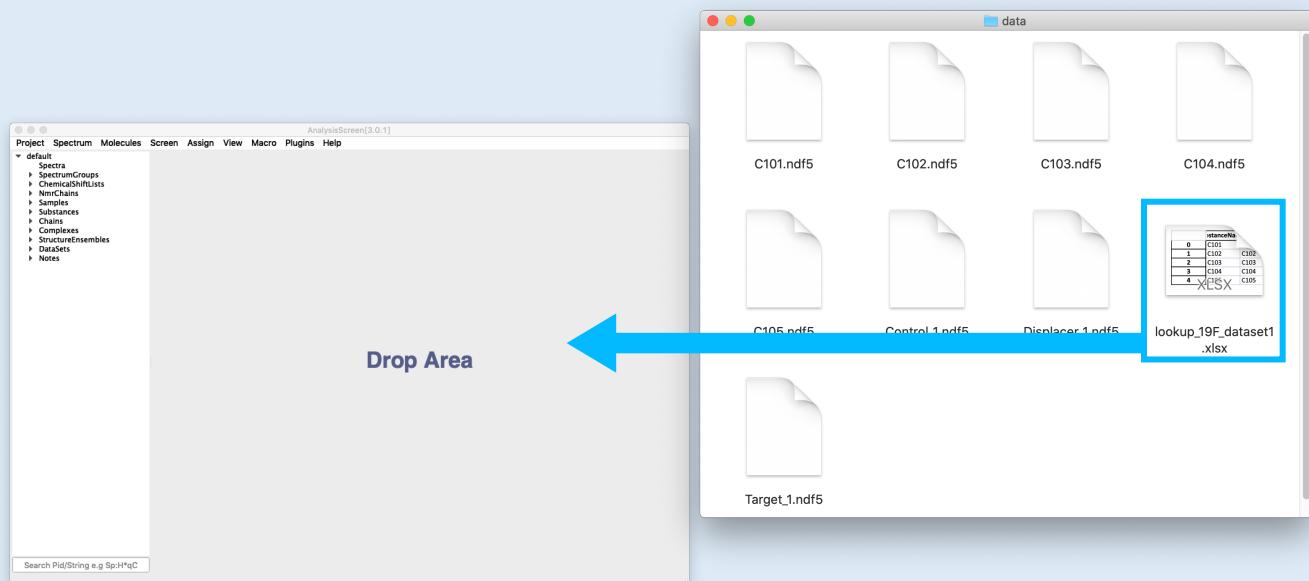
- The other columns record a sample's chemical properties and other information. Fill them in if you want to display them within the software.

To add extra spectra for the same sample, repeat points 1 to 3 as shown in the figure. There is no need to duplicate the samples properties (yellow columns) as long as the sample name is the same. If you add the same information twice, only the first entry will be considered.

To add an additional sample, simply fill in further rows, e.g.:

N	sampleName	spectrumGroupName	spectrumPath	experimentType	sampleComponents	pH	io
2	Sample1	Control		STD.H	component1,component2,component3,component4,component5	5.5	
3	Sample1	Displacer		STD.H			
4	Sample1	OffResonance	STDs/sample1OffReson	STD.H			
5	Sample1	OnResonance	STDs/sample1OnReson	STD.H			
6	Sample2	Control		STD.H	component6,component7,component8,component9,component10	5.5	
7	Sample2	Displacer		STD.H			
8	Sample2	OffResonance	STDs/sample1OffReson	STD.H			
9	Sample2	OnResonance	STDs/sample1OnReson	STD.H			
10	Sample2	STDTarget		STD.H			
11	Sample3	Control		STD.H	component11,component12,component13,component14,component15	5.5	
12	Sample3	Displacer		STD.H			

Import Data from Excel



1F Import Excel Lookup File into AnalysisScreen

- Drag and Drop either your newly created Excel file or the **lookup_19F_dataset1.xls** file located in the **ScreenTutorial/19F/dataset1/data/** directory of the tutorial data from your file browser into the **sidebar** or **drop area** of AnalysisScreen.

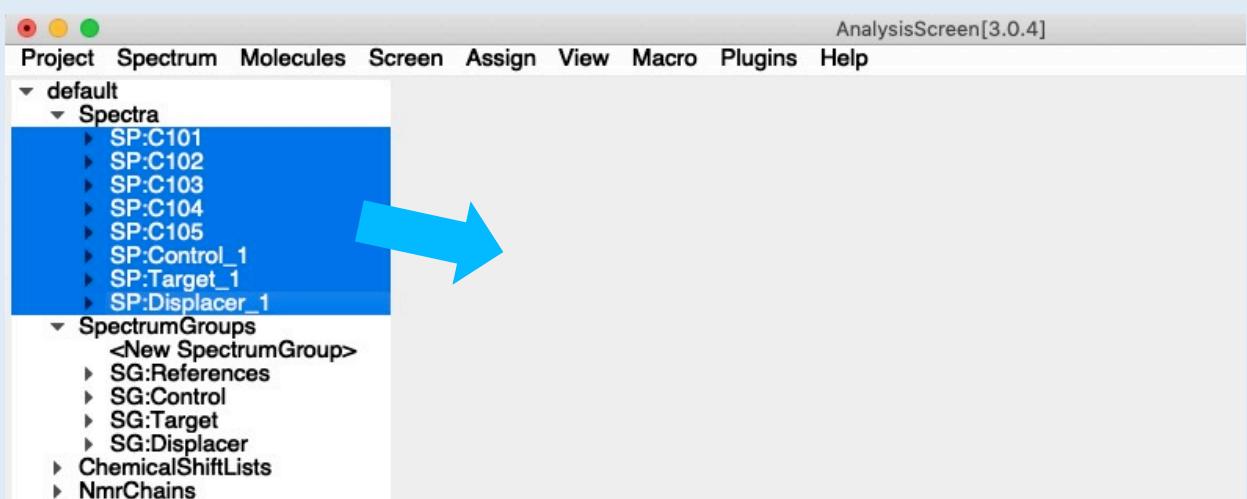
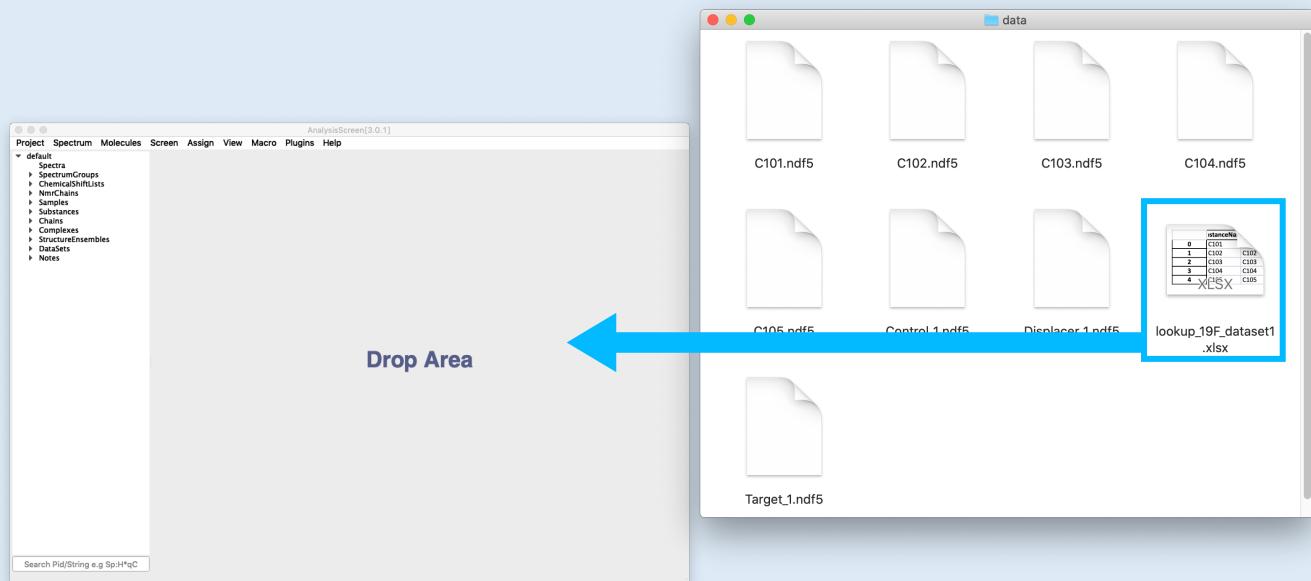
You will now be able to see all the imported data in the sidebar:

- ▼ default
 - ▼ Spectra
 - SP:C101
 - SP:C102
 - SP:C103
 - SP:C104
 - SP:C105
 - SP:Control_1
 - SP:Target_1
 - SP:Displacer_1
 - ▼ SpectrumGroups
 - <New SpectrumGroup>
 - SG:References
 - SG:Control
 - SG:Target
 - SG:Displacer
 - ChemicalShiftLists
 - NmrChains
- ▼ Samples
 - <New Sample>
 - SA:Control_1
 - SA:Target_1
 - SA:Displacer_1
- ▼ Substances
 - <New Substance>
 - SU:C101.
 - SU:C102.
 - SU:C103.
 - SU:C104.
 - SU:C105.
 - Chains
 - Complexes
 - StructureEnsembles
 - StructureData
 - DataTables
 - Collections
 - Notes

Please note that you cannot drop the same lookup file containing the same values into the same project twice. This is because the project cannot create new objects with pre-existing names. When dropping the same file onto a project twice, only the first entries will be used.

You can now proceed to match your peaks in **Section 3**.

2 Manual Peak Matching from Reference Singlets



2A Add your data to the program with an Excel file

If you haven't already added the data from dataset_1 to your project in **Section 1**, then:

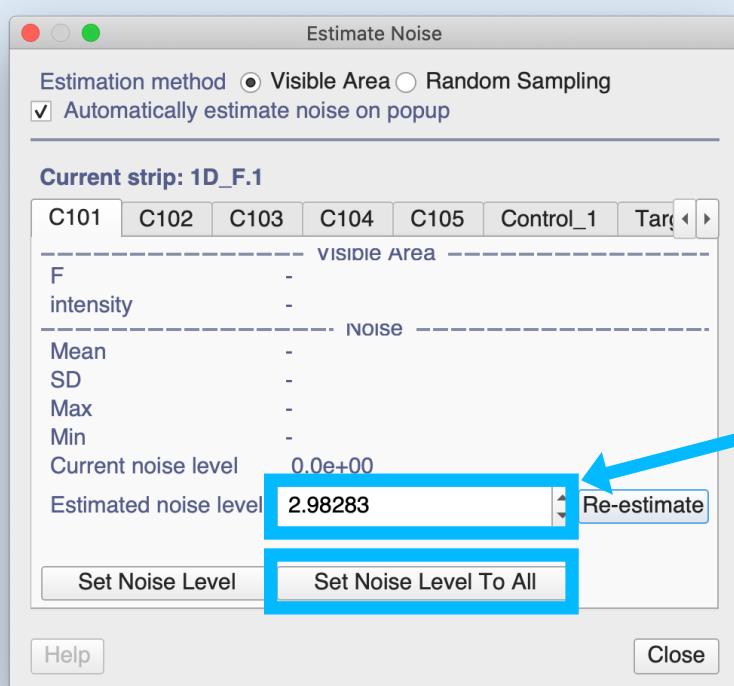
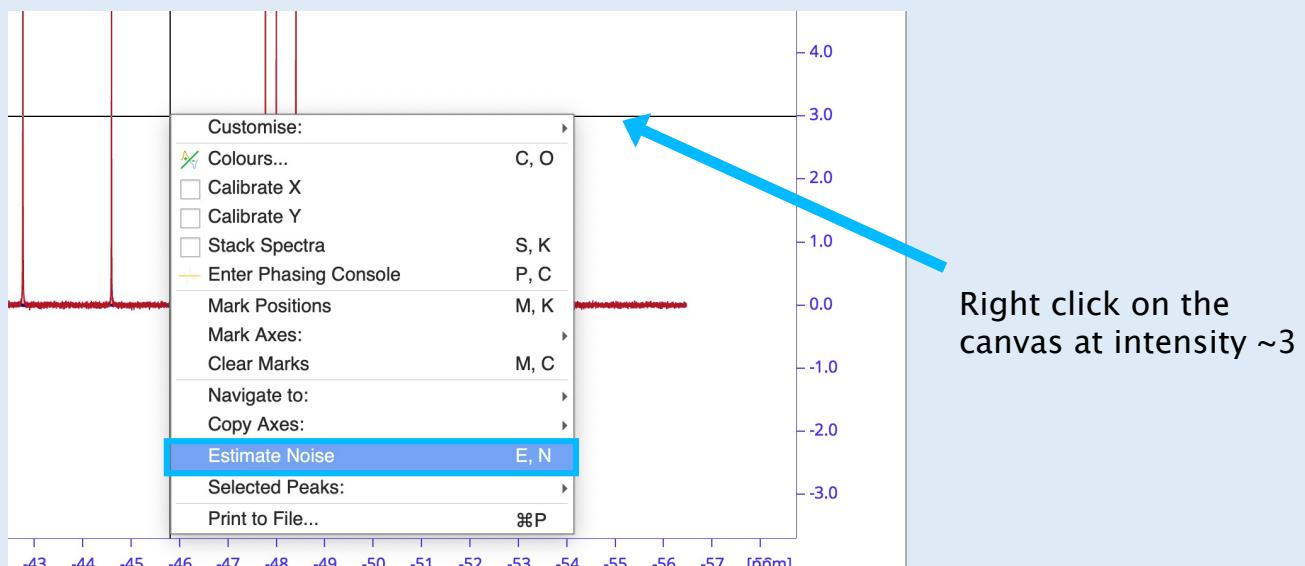
- Drag and drop the Excel file **lookup_19F_dataset1.xlsx** in the **ScreenTutorial/19F/dataset_1/data/** directory into the program.

By loading a correctly formatted Excel file, all the necessary links are automatically established ensuring optimal functioning of the screening tools. See the “HowTos_SidebarObjects” and the “HowTos_ImportDataFromExcel” manuals for more information.

2B Open all spectra

- On the sidebar, expand the **Spectra** branch
- select all spectra and then drag and drop them onto the Drop Area.

Manual Peak Matching from Reference Singlets

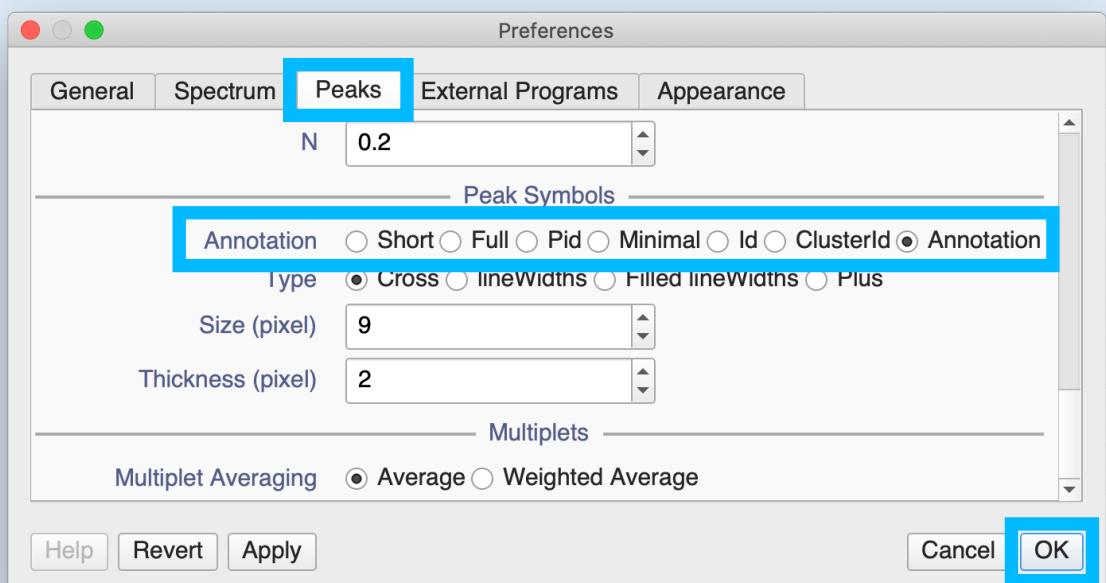
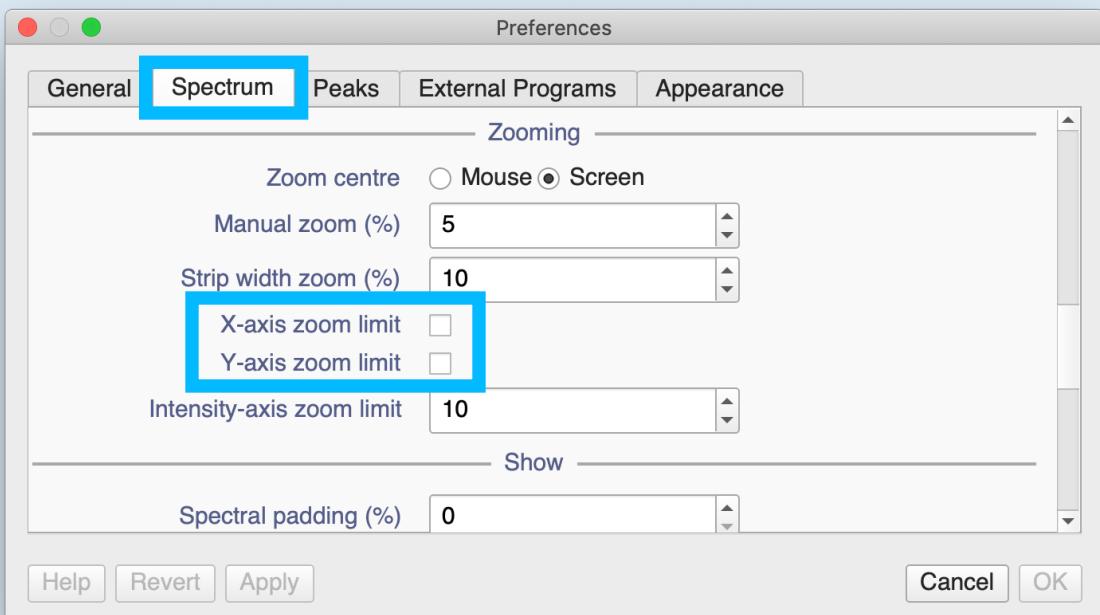


2c Set noise level

In the Spectrum Display:

- place the mouse cursor at a position where you want to set the noise level, for example, at ~ 3 on the Intensity axis.
- right click -> **Estimate Noise**
- in the pop-up click on **Set Noise Level To All**

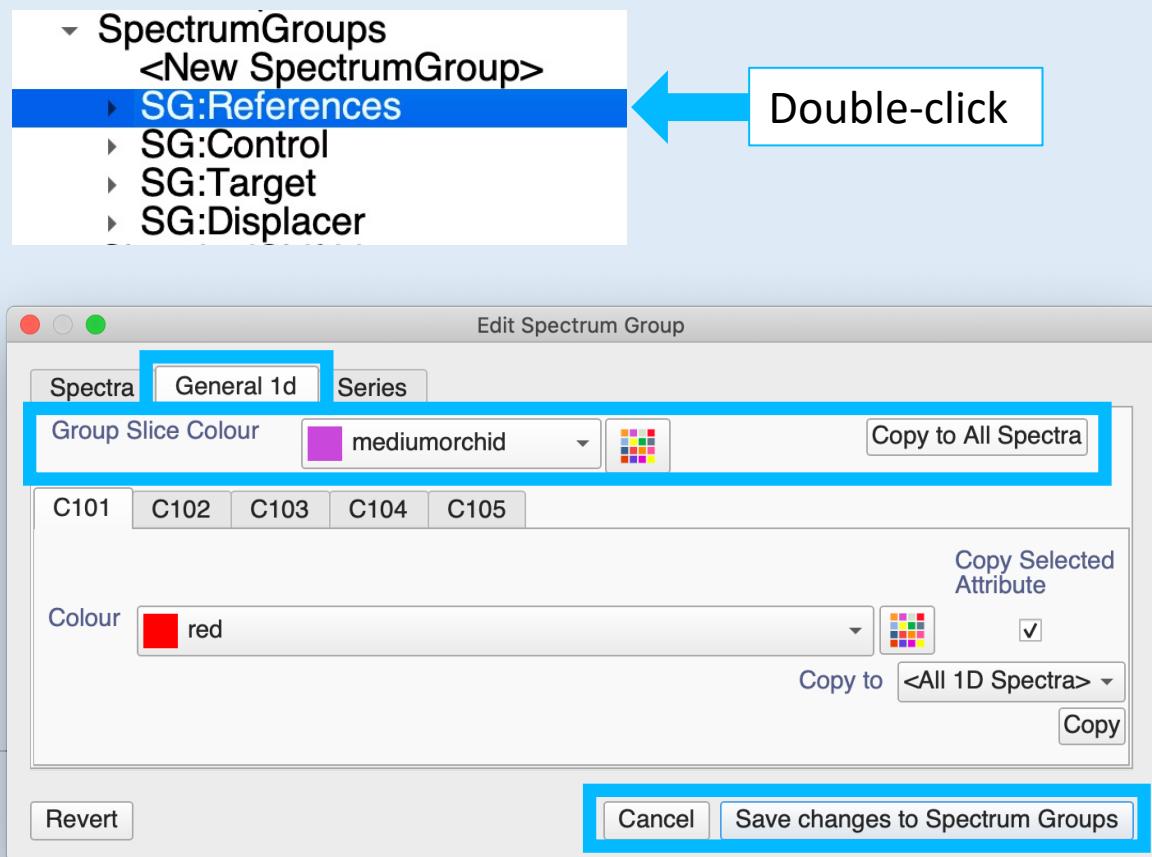
Manual Peak Matching from Reference Singletons



2D Adjust Preferences

- Go to Main Menu -> File -> Preferences
- In the **Spectrum** tab, untick **Apply zoom limit** both for X and Y axes
- In the **Peaks** tab, set **Label** to **Annotation**.

2 Manual Peak Matching from Reference Singletons



2E Colour spectra by Spectrum Group

If you like, you can set all the spectra in one Spectrum Group to be displayed in the same colour in one go. It can be helpful to colour your Reference / Control / Target / Displacer spectra in the same colours in each project that you can quickly recognise which group a spectrum belongs to.

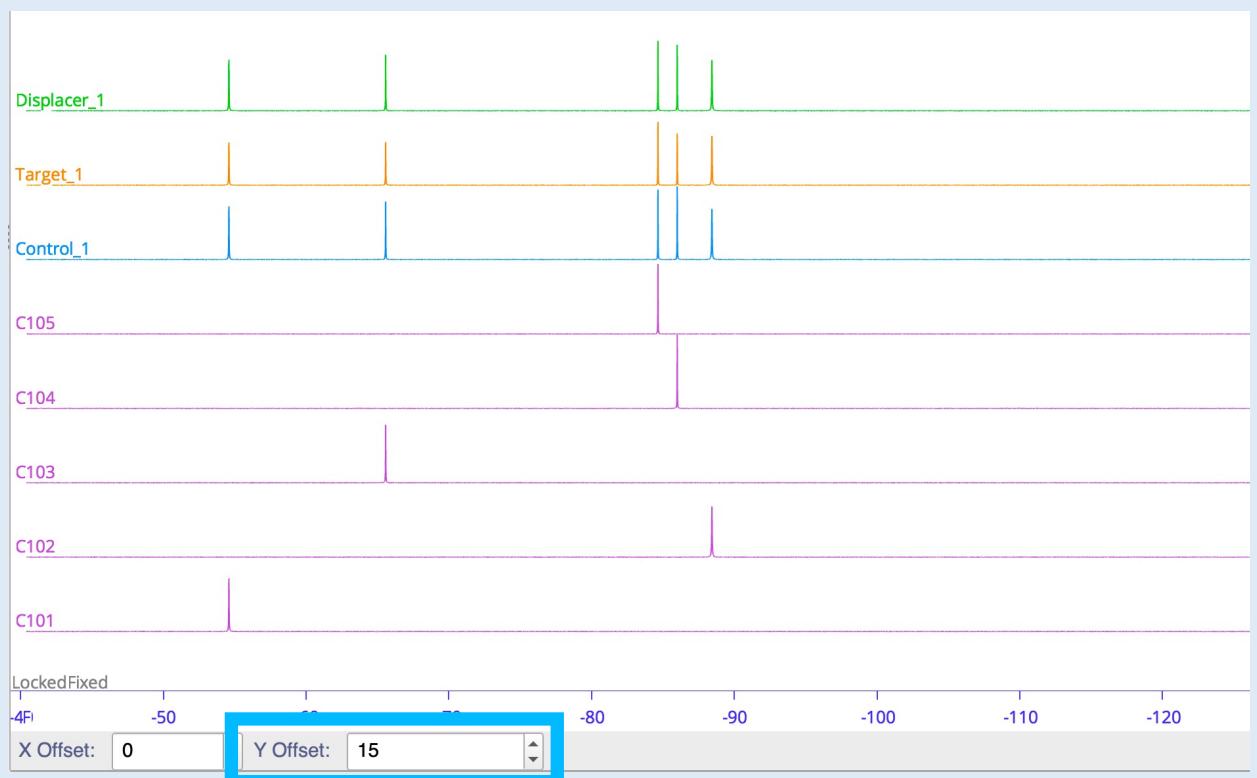
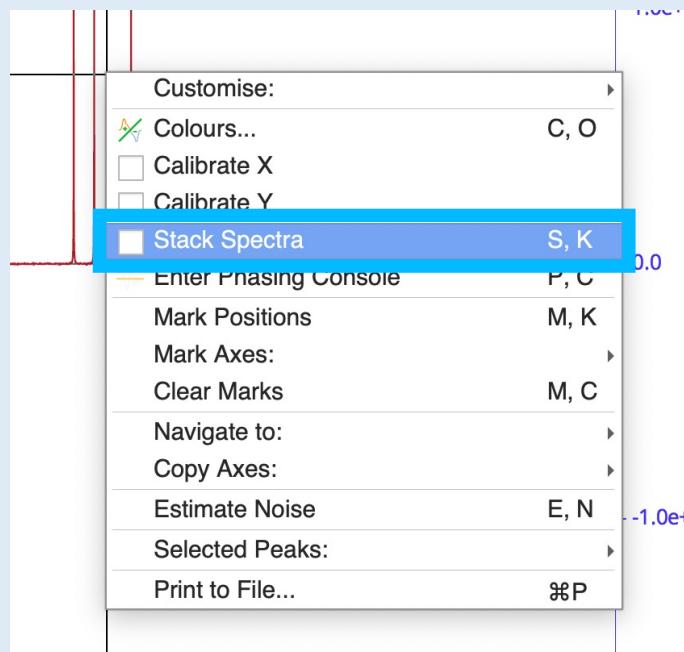
- Double-click on the **SG:References** Spectrum Group in the sidebar and go to the **General 1d** tab.
- Select a **Group Slice Colour** and click on **Copy to All Spectra**.
- Click on **Cancel** or **Save Changes to Spectrum Groups** to close the pop-up.

Throughout this tutorial we will use the following colour coding:

References	pink (mediumorchid)
Control	blue (dodgerblue)
Target	orange (darkorange)
Displacer	green (limegreen)

Change the colouring of your spectrum groups to reflect this colouring if you like.

Manual Peak Matching from Reference Singlets

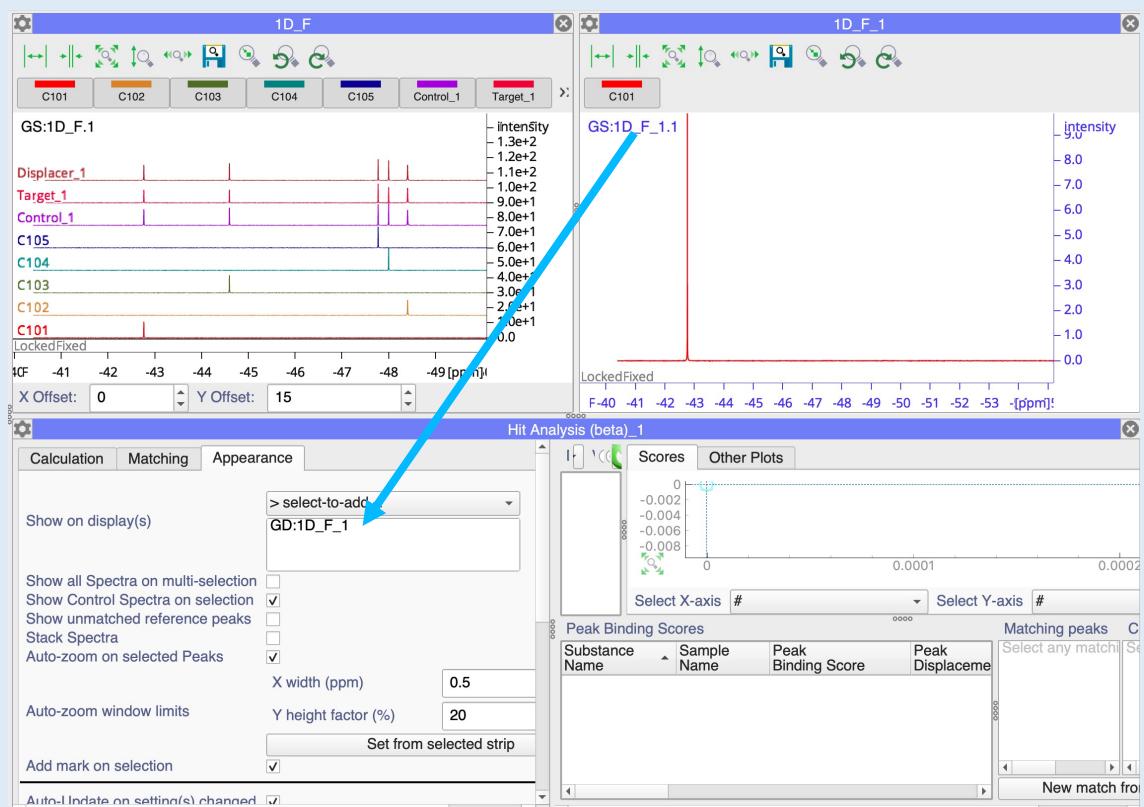
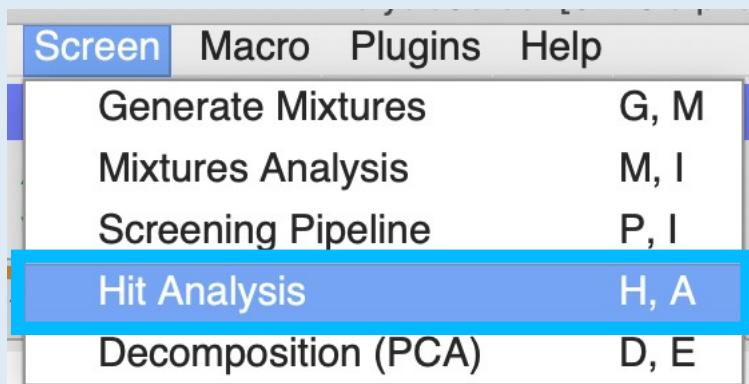


2F Stack Spectra

In the Spectrum Display :

- right click -> **Stack Spectra** (shortcut **SK**)
- Y Offset: 15**
- Zoom in/out as required to show all spectra in the display
- Scroll the mouse-wheel over the Intensity axis to adjust the Y-range

Manual Peak Matching from Reference Singletons



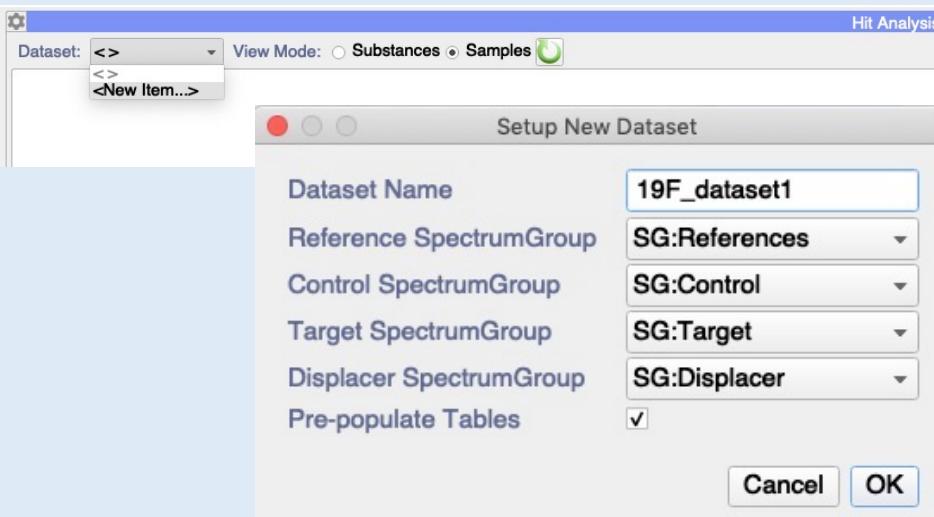
2G Open the Hit Analysis Module

- Go to Main Menu → Screen → Hit Analysis (shortcut HA)

2H Open a second Spectrum Display and re-arrange layout

- Open any spectrum next to the already opened stacked display
- Click on the Hit Analysis module **Settings** gear icon :
- **Appearance tab**
- **Show on display(S) -> right click -> Remove All**
- Add the new unstacked **Spectrum Display** (e.g. GD:1D_F_1) to the list
The module will soon perform a series of dynamic actions in this display.
- Close the settings panel.
- Re-arrange the Hit Analysis Module below the Spectrum Displays to show all its widgets on the screen.

Manual Peak Matching from Reference Singletons



Hit Analysis (beta):1								
#	Substance Name	Substance Binding Score	Substance Displacement Score	Substance Matching Score	Substance Label	Relative S/N Ratio	Sample Name	Target Name
1	C101	0.0000	0.0000				Target_1	
2	C102	0.0000	0.0000				Target_1	
3	C103	0.0000	0.0000				Target_1	
4	C104	0.0000	0.0000				Target_1	
5	C105	0.0000	0.0000				Target_1	

The **Hit Analysis** module contains two views: by **Substance** and by **Sample**. This selection determines the behaviour of the two main tables.

The top table contains a list of the substances or samples depending on the selected view. The lower table contains a list of all the peaks for the single substance reference spectrum or for all substances present in a selected sample.

2 Create a new Screening Dataset

In the Hit Analysis module:

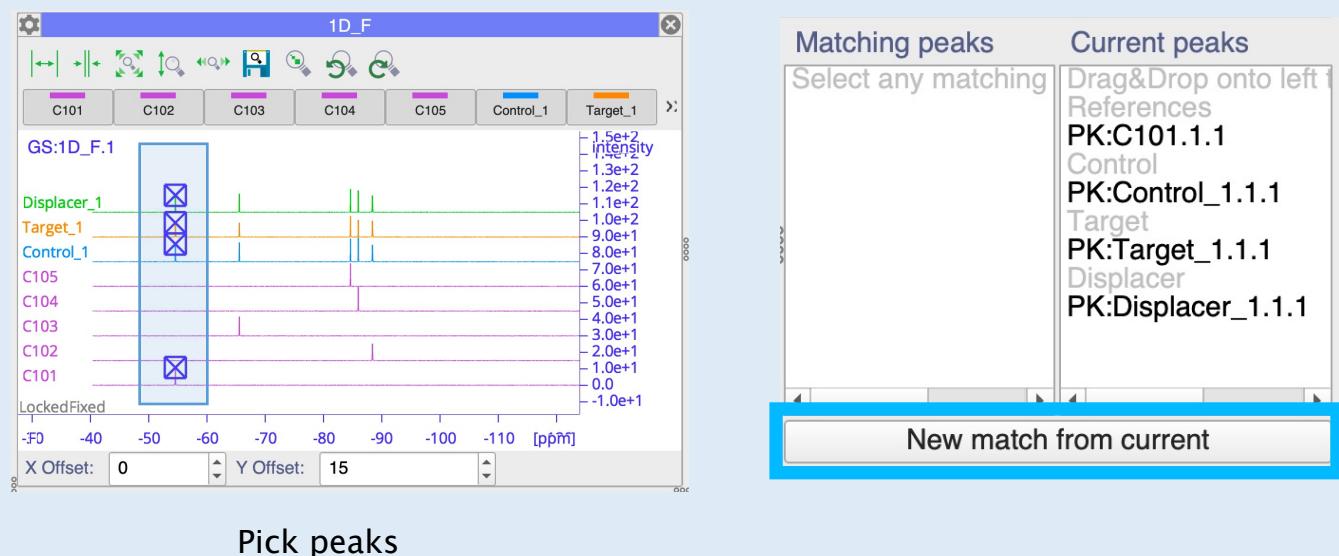
- Select <New Item...> in the **Dataset** pulldown,
- Change the name to **19F_dataset1** or keep the default.
- Press **Ok** to proceed and close the popup.

All substances and samples in the project are now organised in the tables and ready to create **Peak Matches** and scores.

- Set View Mode as **Substances**.
- In the Substance Table, select the first entry, e.g., **C101**.

The Spectrum Display **GS:1D_F_1.1** will update to show the Control–Target–Displacer and C101 spectra.

Manual Peak Matching from Reference Singlets



Pick peaks

2K Pick and match peaks

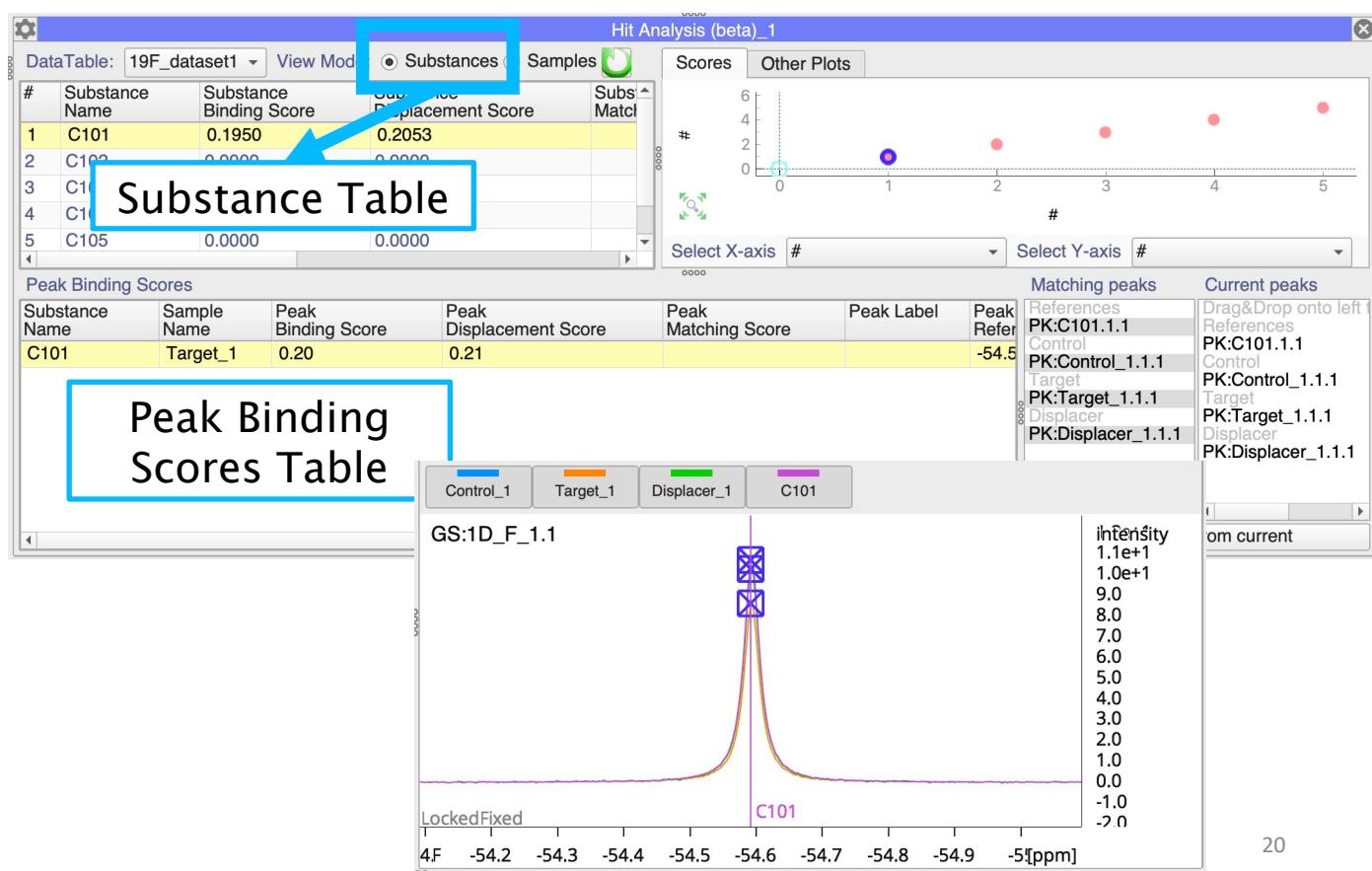
Pick the peaks at ~ -54.59 ppm in the stacked Spectrum Display:

- Use **CTRL** (or **CMD** for Mac) + **SHIFT** + **Left-drag** to create a (blue) picking region to include all spectral signals around -54 ppm.
- This signals correspond to the Substance **C101**.

In the Hit Module:

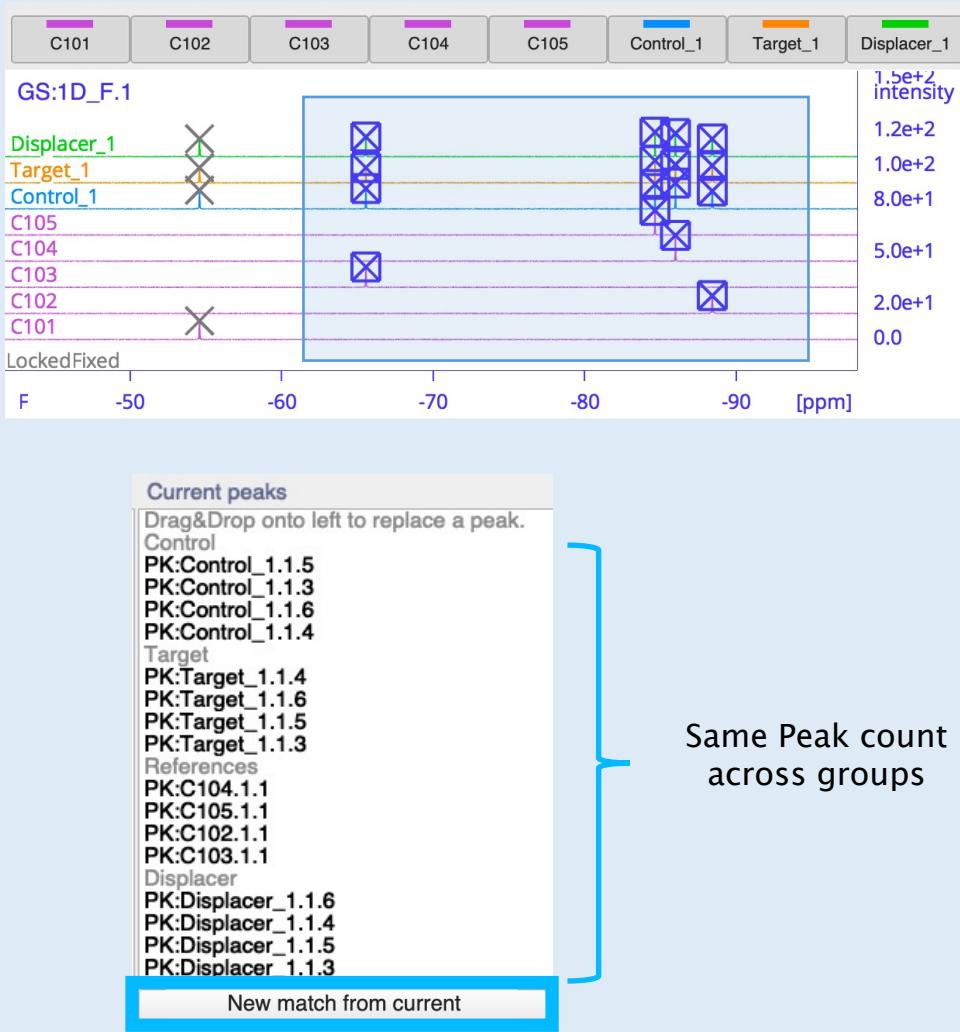
- the **Current peaks** will be shown in the bottom-right hand corner
- Press **New match from current** to create a new **Peak Match**

A new row will appear for the Substance **C101** in the **Peak Binding Scores** table which corresponds to that peak match. Select the row to navigate to the corresponding peaks in the spectrum display.



2 Manual Peak Matching from Reference Singletons

Multi-picking and semi-automatic peak matching



2L Semi-automatic peak picking and matching

In the Spectrum Display:

- Pick the remaining spectral signals by creating a larger picking region with the shortcut **CTRL** (or **CMD** for Mac) + **SHIFT** + **Left-drag**

In the Hit Analysis Module:

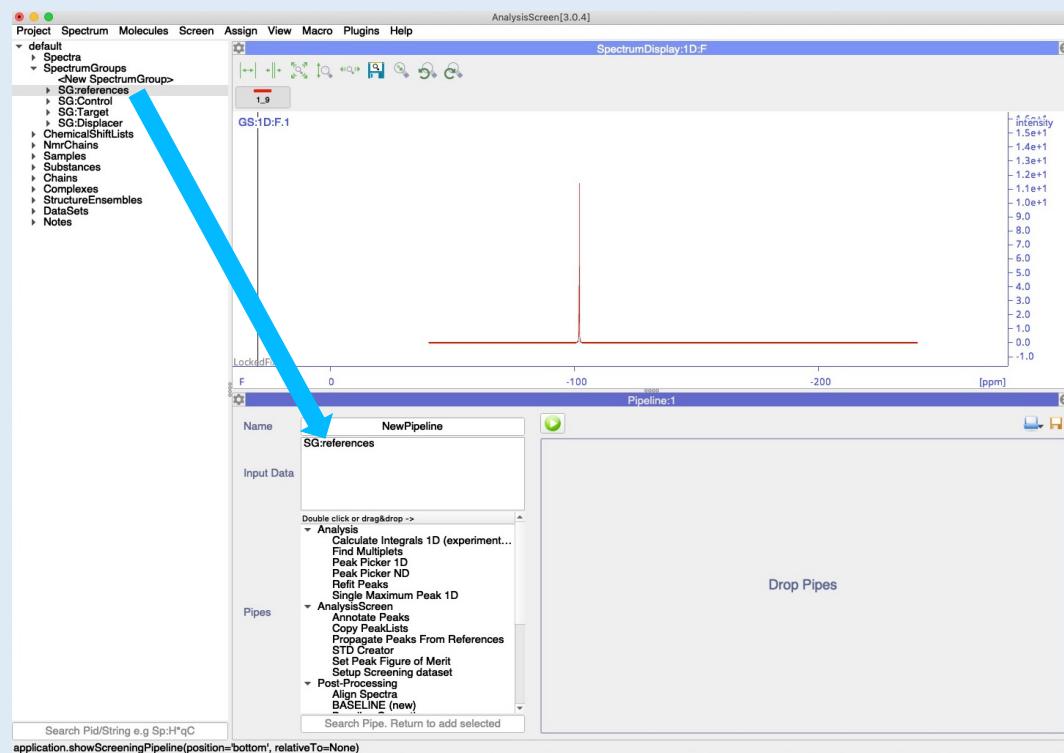
- The **Current peaks** will appear in the bottom-right hand corner. Make sure there is an equal number in each category, e.g.: 4 Reference peaks, 4 Control Peaks etc... in any order.
- Press **New match from current**.

The tables will update:
for each substance you
select in the upper table, a
peak binding score is visible
in the lower table.

Dataset: 19F_dataset1		View Mode: <input checked="" type="radio"/> Substances <input type="radio"/> Samples	
#	Substance Name	Substance Binding Score	Substance Displacement Score
4	C104	0.2952	0.3392
3	C103	0.2595	0.1343
1	C101	0.1950	0.2053
5	C105	0.0951	0.0002
2	C102	0.0301	-0.0000

Peak Binding Scores		
Substance Name	Peak Binding Score	Peak Displacement Score
C104	0.2952	0.3392

Automatic Peak Matching from Reference Singletons



**shortcut
PI**

Automatic peak matching and binding analysis

In this section, a pipeline is built to inspect changes between spectra recorded for mixtures of substances with and without a target, and with target and displacer.

A completed project is provided in the dataset2 folder.

3A Open a new dataset

- Open a new project with **Main Menu** → **File** → **New Project**
 - Drag & drop the **lookup_19F_dataset2.xlsx** file from the **ScreenTutorial/19F/dataset_2/Data** folder into the sidebar or drop area.
- See the **Section 1** for more information about setting up Excel files.

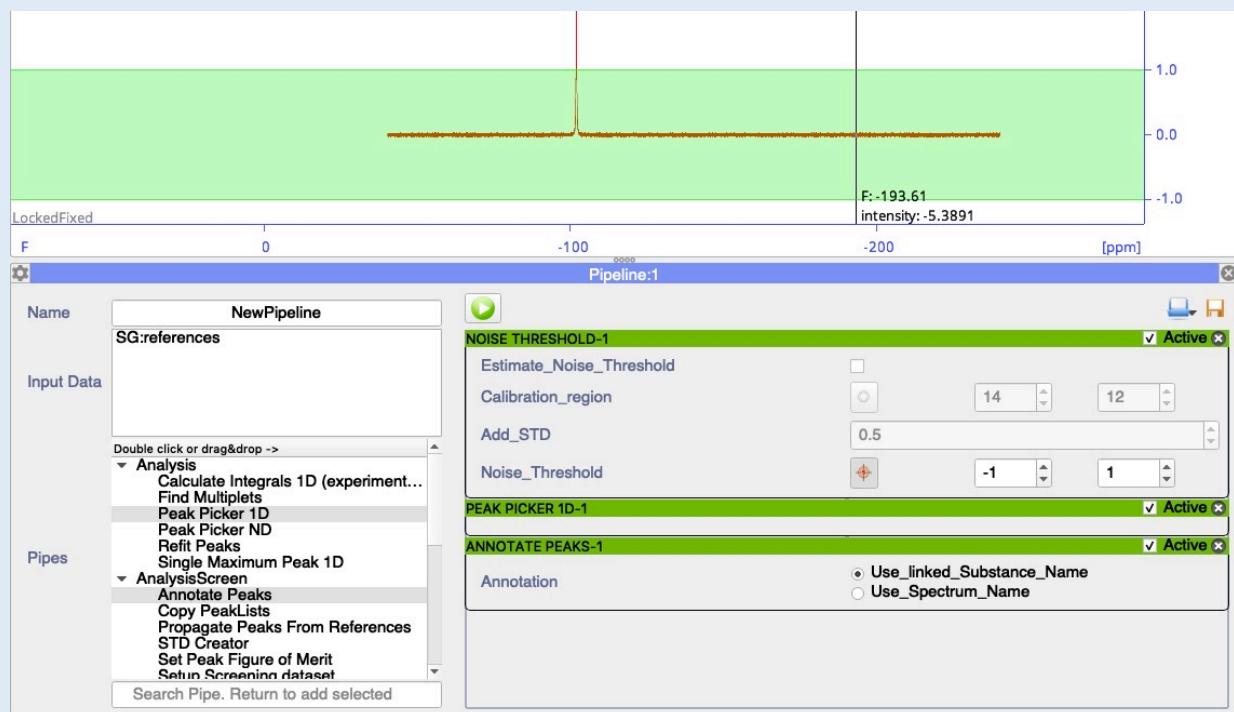
3B Open the first reference spectrum

- select the first spectrum in the Sidebar, drag and drop it onto the Drop Area.
- If you like, colour your spectra by their Spectrum Group type as shown in **Section 3E**.

3C Open the pipeline module, shortcut PI

- Open the pipeline module from the main menu:
Menu → **Screen** → **Pipeline**
- or use the **shortcut PI**
- Expand the Sidebar branch for **SpectrumGroups**
- Select **SG:References**, drag & drop it into the **Input Data** of the Pipeline

Automatic Peak Matching from Reference Singletons



Multiple Pipes can be added and re-ordered by holding and dragging the green top bar. See [HowTos_Pipelines](#) for more information.

3D Pick peaks in Reference Singleton spectra

- In the Pipes list widget search for and add the following pipes to the pipeline area by double clicking the pipe name or via drag & drop:

1. Generic > Noise Threshold

- untick **Estimate Noise threshold**
- Click on the Target button
- Insert the **Noise Threshold** values -1, 1 either by dragging the green lines that appears on displays or inserting the values in the entries

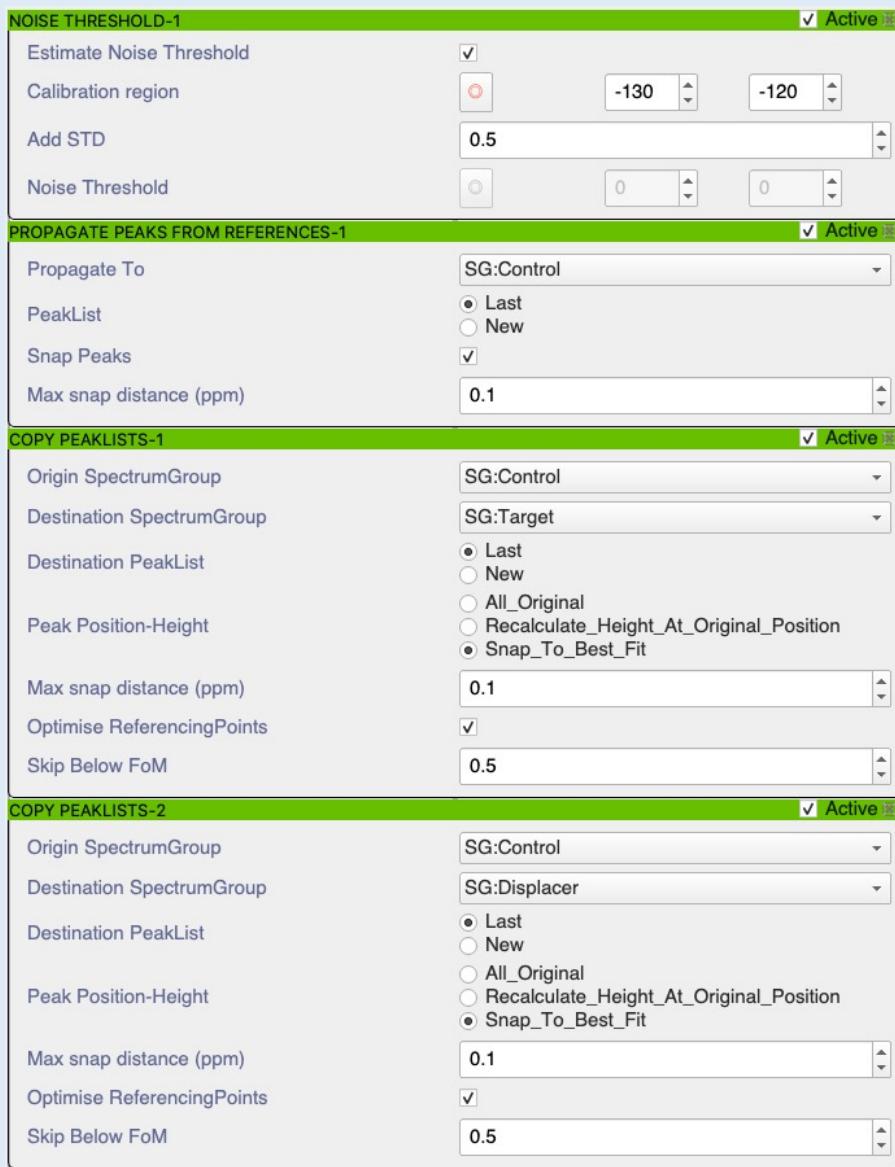
2. Analysis > Peak Picker 1D

3. AnalysisScreen > Annotate Peaks

- Tick **Use linked Substance Name**

- Run the pipeline using the green play button (click once only!). A popup will appear when completed.
- If you wish, you can save this or any other pipeline you create by clicking on the icon. You can open a pipeline with but check your parameters are set correctly, as not all are saved.

Automatic Peak Matching from Reference Singletons



Propagate Peaks will copy peaks from several (Reference Singleton) peak lists and collate them into a smaller number of (Control mixture) peak lists.

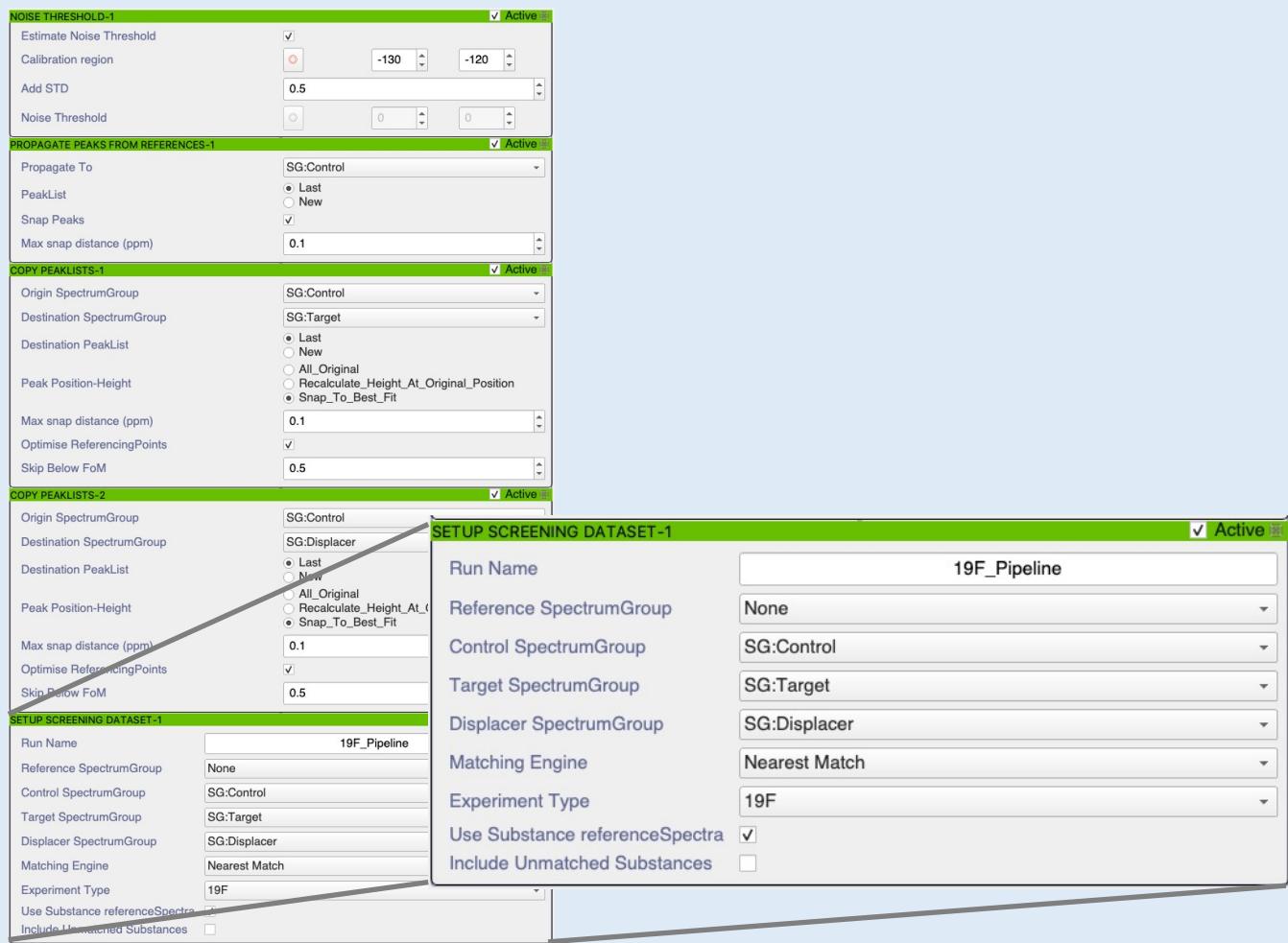
Copy Peaks will copy the peak lists one to one (from the Control mixtures to the Target and Displacer mixtures).

3E Setup screening pipeline

- Clear the input data (**right-click** → **Clear all**)
- On sidebar, multiselect 'SG:Control', 'SG:Target', 'SG:Displacer' and then drag & drop into the pipeline **Input Data** area.
- Close all pipes (**right-click** on any pipe header → **Close All**)
- In the list of **Pipes** search for and add these pipes to the pipeline area:
 - 1. Generic > Noise Threshold**
 - Calibration region: -120, -130 ppm
 - 2. AnalysisScreen > Propagate Peaks from References**
 - Propagate to: SG:Control
 - 3. AnalysisScreen > Copy PeakLists (1)**
 - Origin SpectrumGroup: SG:Control
 - Destination SpectrumGroup: SG:Target
 - 4. AnalysisScreen > Copy PeakLists (2)**
 - Origin SpectrumGroup: SG Control
 - Destination SpectrumGroup: SG:Displacer

continued....

Automatic Peak Matching from Reference Singletons



...continued

5. AnalysisScreen > Setup Screening dataset

- **Run name:** 19F_dataset
- **Reference SpectrumGroup:** None
- select **SG:Control**, **SG:Target**, **SG:Displacer** for their respective entries
- **Matching Engine:** Nearest Match
- tick **Use Substance ReferenceSpectra**

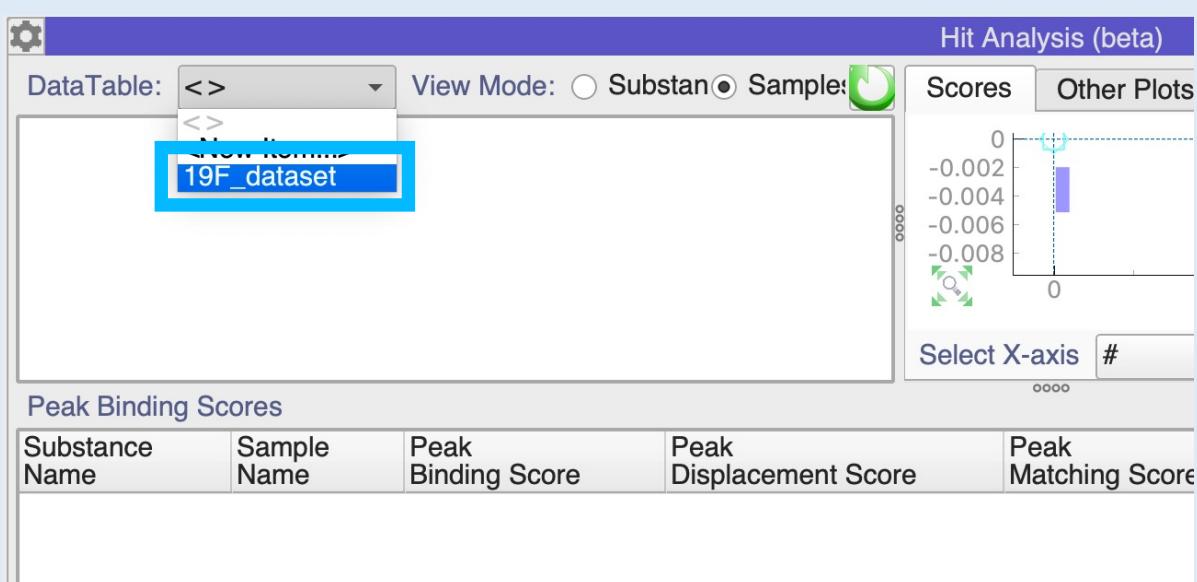
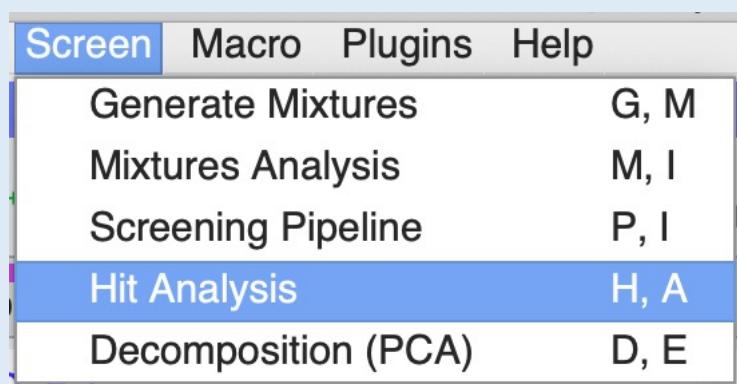
- Run the pipeline using the green play button.

The calculations should take less than a minute. When it is finished you should see a new entry under DataTables in your sidebar:

- ▼ **DataTables**
 - <New DataTable>
 - DT:19F_Pipeline
- ▶ **Collections**
- ▶ **Notes**

- Save the pipeline if you wish.

4 Binding Scores and Calculation Engines



4A Open data in program

You can now either continue with your project from **Section 3** or load a project containing which has already completed those steps:

- Drag and drop the **ScreenTutorial/19F/dataset_2/completed_dataset2 ccpn** folder in the sidebar or Drop Area.
- Open a spectrum in a Spectrum Display module.

Open the Hit Analysis module:

- Go to **Main Menu** → **Screen** → **Hit Analysis** or use shortcut **HA**.
- From the **DataTable** drop-down menu select **19F_dataset** and wait a few moments for the tables to be updated.

4 Binding Scores and Calculation Engines

#	Substance Name	Substance Binding Score	Substance Displacement Score	Substance Matching Sc
1	C101	0.1504	0.3317	5
2	C1010	0.0498	0.9048	5
3	C1011	0.1947	0.0258	5
4	C1012	0.0254	0.9860	5
5	C1013	0.0057	0.0050	5

4B Scoring Engines

In the Hit Analysis Module:

- Select **Substances** in the **View Mode** and sort by **Substance Binding Score** (click on the column header).

The **Substance Binding Score** gives an indication of the spectral changes between the Control and Target spectra for peaks matched to the Substance reference spectrum. Therefore, it can be used to assess the substance binding quality.

Peaks are compared by one of the following properties: height, linewidth or volume. (Note that linewidths and volumes are not calculated by default when peaks are picked and have to be determined with **Estimate Volumes**.)

Open the Hit Analysis Module settings from the gear icon and select the **Calculation** tab.

The binding score is given by the calculation **Engine**.

- Change the default by selecting one of the options. When hovering over the labels for each option, a Tooltip window will show the equation used.
- You can also define your own equation in the free entry box:

use **V1** and **V2** to define the variables for the calculation matrix

- **V1** represents each Ligand signal in the presence of the target (Target)
- **V2** represents each Ligand signal in the absence of the target (Control)

The following arithmetic operations are supported:

$\sim + \sim$, $\sim - \sim$, $\sim * \sim$, \sim / \sim , $\sim * \sim$, $\sim \% \sim$, \sim / \sim

The same applies to the **Displacement Engine** which gives the **Substance Displacement Score**. When writing your own equation, define the Displacer signal with the variable **V3**.

continued...

4 Binding Scores and Calculation Engines

Calculation Mode: Matching

Engine: AbsoluteRelativeChange

Displacement Engine: DisplacementFraction

Total Score: sum

Scale Score (0-1):

Dataset: Matching

#	Substance Name	Substance Binding

Substance Name: Peak Position Reference

Hit Analysis (beta):1

Dataset: 19F_dataset1 View Mode: Substances Samples

Scoring Engines (continued)

Note that the **Substance Binding Score** is derived from all the peaks matched to that substance.

In this ^{19}F demo dataset, only one observation is recorded per reference spectrum (as is typical for ^{19}F data).

However, if multiple peaks per substance are present (as might be the case for ^1H data), the total **Substance Binding Score** will be given by one of :

min, max, std, mean or sum

of all the single peak binding scores.

- Select your preferred option from the **Total Score** Settings.

If the **View Mode** is set to **Samples**, a **Sample Binding Score** is shown and derived from the **Substance Binding Scores** of all substances in that sample.

As soon as you change settings or modify/fit peaks, the whole module will update and recalculate all scores automatically. For larger datasets this can be time-expensive. You can disable this feature:

- In **Settings -> Appearance tab**
- untick **Auto-Updates on setting(s) changed**
- untick **Auto-Updates on peak(s) changed**

The refresh button will turn red whenever changes are detected while working on the dataset. Click the refresh button to update all scores after which it will turn green.

4 Binding Scores and Calculation Engines

Hit Analysis Module Table Selections

The Hit Analysis module has multiple dynamic selections

Selecting a row on the **Substances table** will:

1. List all contributing peak matches in the **Peak Binding Scores** table
2. Display all the spectra associated to the binding match
3. Select the relevant item in the Scores Scatter plot on the right

Selecting a row on the **Samples table** will:

1. List all contributing peak matches in the **Peak Binding Scores** table for all the substances present in the sample
2. Display all the spectra associated to the binding match, including all reference spectra

Selecting a row on the **Peak Binding Scores table** will:

1. Select all peaks included in the match
2. Navigate to the Peak Position
3. Populate the **Matching/Current Peaks** lists widgets

Double-clicks on tables will re-execute the single selection.

The screenshot shows the Hit Analysis module interface. At the top, there is a navigation bar with a gear icon, a dataset dropdown set to "19F_dataset1", and a "View Mode" radio button group where "Substances" is selected. Below this is a table titled "Substances Table". The table has columns: #, Substance Name, Substance Binding Score, Substance Displacement Score, Substance Matching Score, Substance Label, Relative S/N Ratio, and Sample Name. Rows 1 through 4 are shown, with row 1 highlighted in yellow. An orange box highlights the first two columns of the table. A blue arrow points from the text "Substances Table" to the highlighted area. Below the table is a section titled "Peak Binding Scores" containing a smaller table with columns: Substance Name, Peak Binding Score, Peak Displa, and Peak Position Displacer. Rows for C101 and C102 are listed. A blue box highlights the first three columns of this table. A blue arrow points from the text "Peak Binding Scores Table" to the highlighted area.

#	Substance Name	Substance Binding Score	Substance Displacement Score	Substance Matching Score	Substance Label	Relative S/N Ratio	Sample Name
1	C101	0.0010	0.5321				Target_1
2	C102						Target_1
3	C103						Target_1
4	C104						Target_1
5	C105	0.0005	0.0000				Target_1

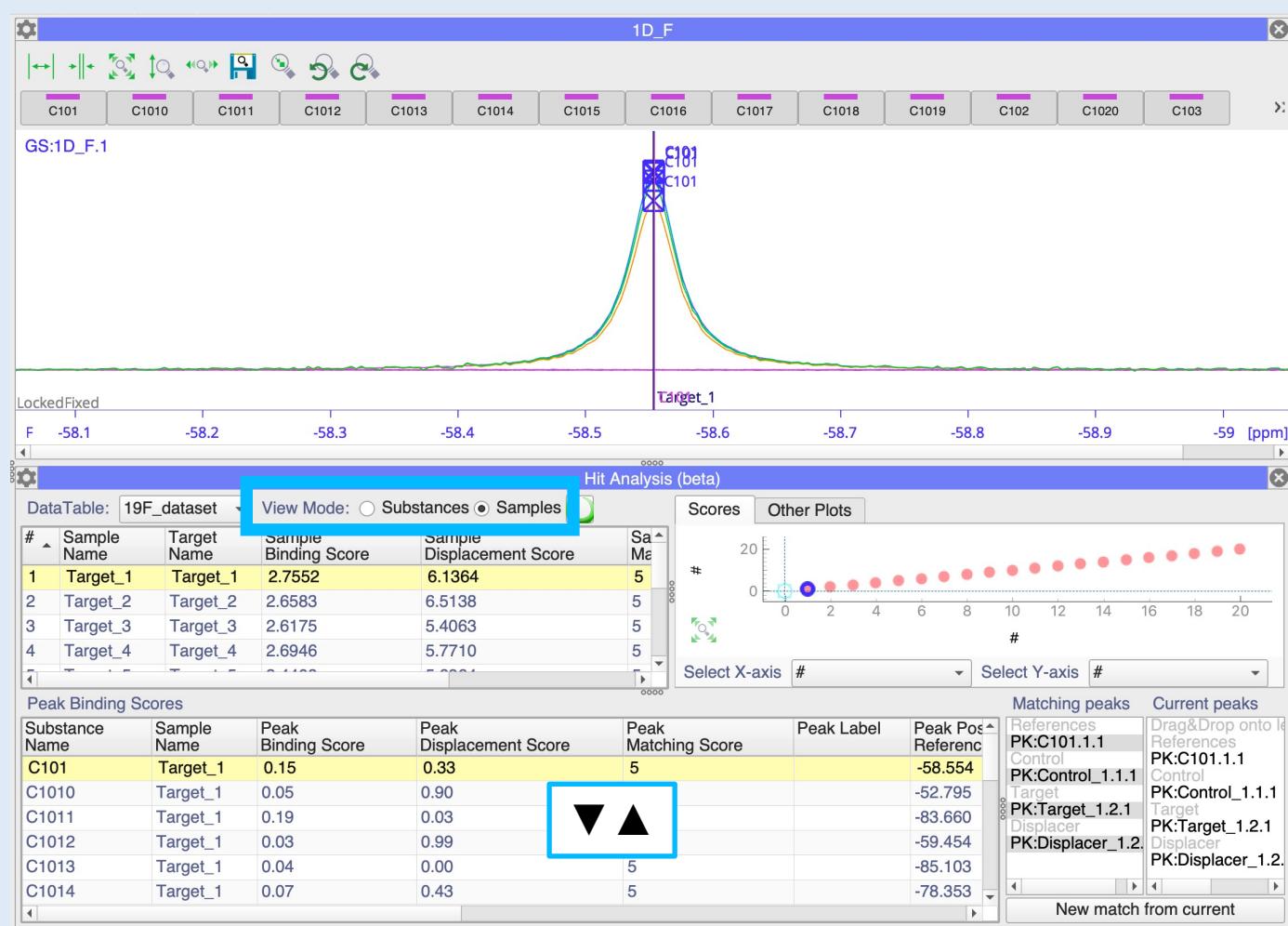
Substance Name	Peak Binding Score	Peak Displa	Peak Position Displacer
C101	0.0010	0.5321	-54.5931
C102	-0.0082	0.3033	-88.4319

The screenshot shows the Hit Analysis module interface. At the top, there is a navigation bar with a gear icon, a dataset dropdown set to "19F_dataset1", and a "View Mode" radio button group where "Samples" is selected. Below this is a table titled "Samples Table". The table has columns: #, Sample Name, Target Name, Sample Binding Score, Sample Displacement Score, Sample Matching Score, and Comment. Row 1 is highlighted in yellow. A blue box highlights the first four columns of the table. A blue arrow points from the text "Samples Table" to the highlighted area. Below the table is a section titled "Peak Binding Scores" containing a smaller table with columns: Substance Name, Peak Binding Score, Peak Displacement Score, Peak Position Reference, Peak Position Control, Peak Position Target, and Peak Position Displacer. Rows for C101, C103, C105, C104, and C102 are listed. A blue box highlights the first six columns of this table.

#	Sample Name	Target Name	Sample Binding Score	Sample Displacement Score	Sample Matching Score	Comment
1	Target_1	Target_1				

Substance Name	Peak Binding Score	Peak Displacement Score	Peak Position Reference	Peak Position Control	Peak Position Target	Peak Position Displacer
C101		0.2053	-54.5932	-54.5931	-54.5931	-54.5931
C103		0.1343				65.5728
C105		0.0001				84.6519
C104		0.3391				86.0021
C102		0.0003	-88.4319	-88.4319	-88.4319	-88.4319

4 Binding Scores and Calculation Engines



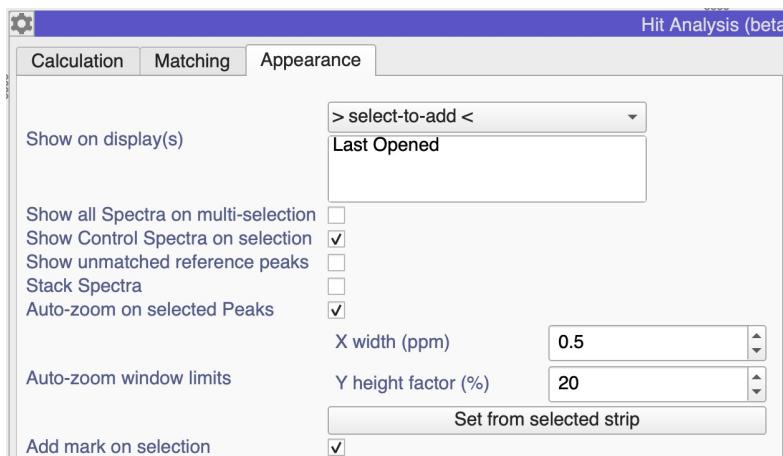
4c Looking through peaks manually

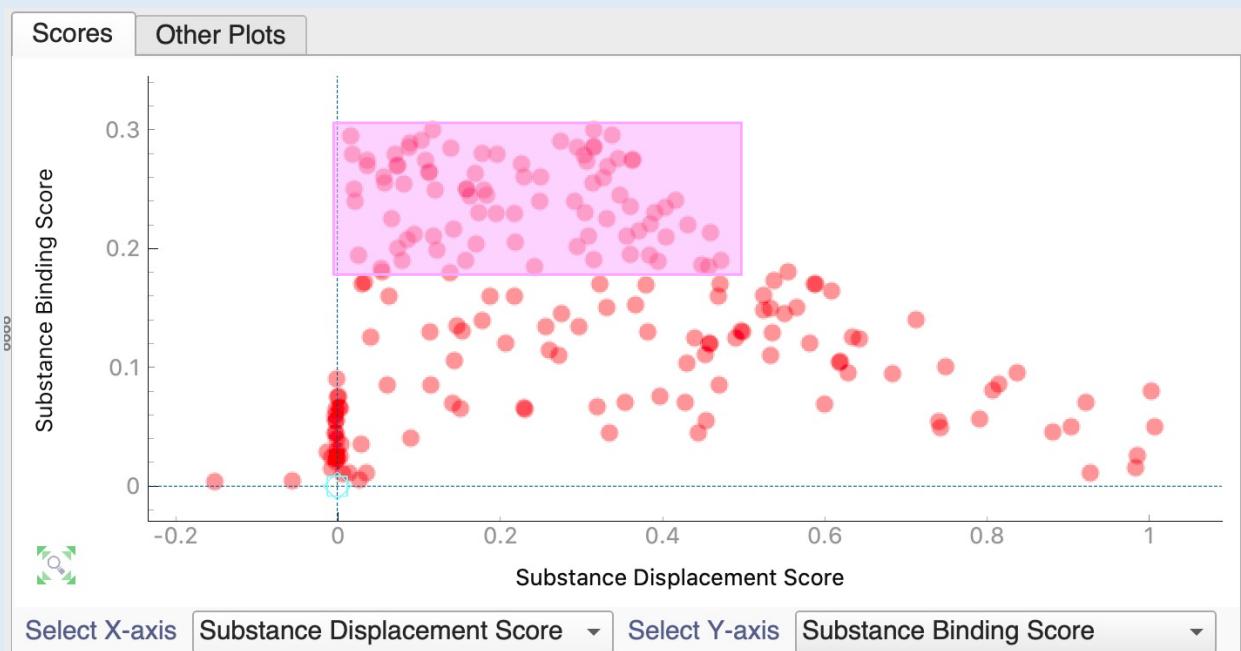
Here is an easy way to scan through all your peaks by hand fairly quickly to see if you spot anything interesting or unusual:

- Set the **View Mode** to **Samples**, so that the upper table in the Hit Analysis module shows one sample per row.
- Move to the **Peak Binding Scores Table** below and use your **up/down arrow keys** to move down the list.

Each time the Spectrum Display will automatically focus on the peak in question.

- Once you have finished going through the peaks in one sample, move to the next sample in the **Samples Table** and go through the next set of peaks.
- In the **Appearance** tab of the Hit Analysis module **Settings** you can change the the auto-zoom options if you would prefer these to be different.





5A Plots

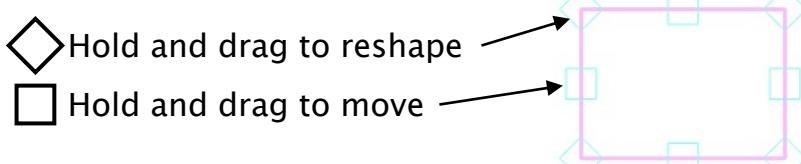
The Hit Analysis module has a Plots window in which you can select to plot any two variables of your choice against one another. If you have done automatic peak matching, you could, for example, plot the **Substance Matching Score** along the y-axis in order to find outliers quickly.

- Selecting an item in a table will select it in the plot and vice versa.

You can manipulate the plot and its points like a spectrum and peaks:

- Zoom with the **mouse wheel**, either on the plot or selectively on a single axis
- Move the plot around with **left-drag**
- (Multi-)select items with **Ctrl/Cmd + left-click**
- Select items in an area with **Ctrl/Cmd + left-drag**

This will draw a Region of Interest (ROI) box which you can change as follows:

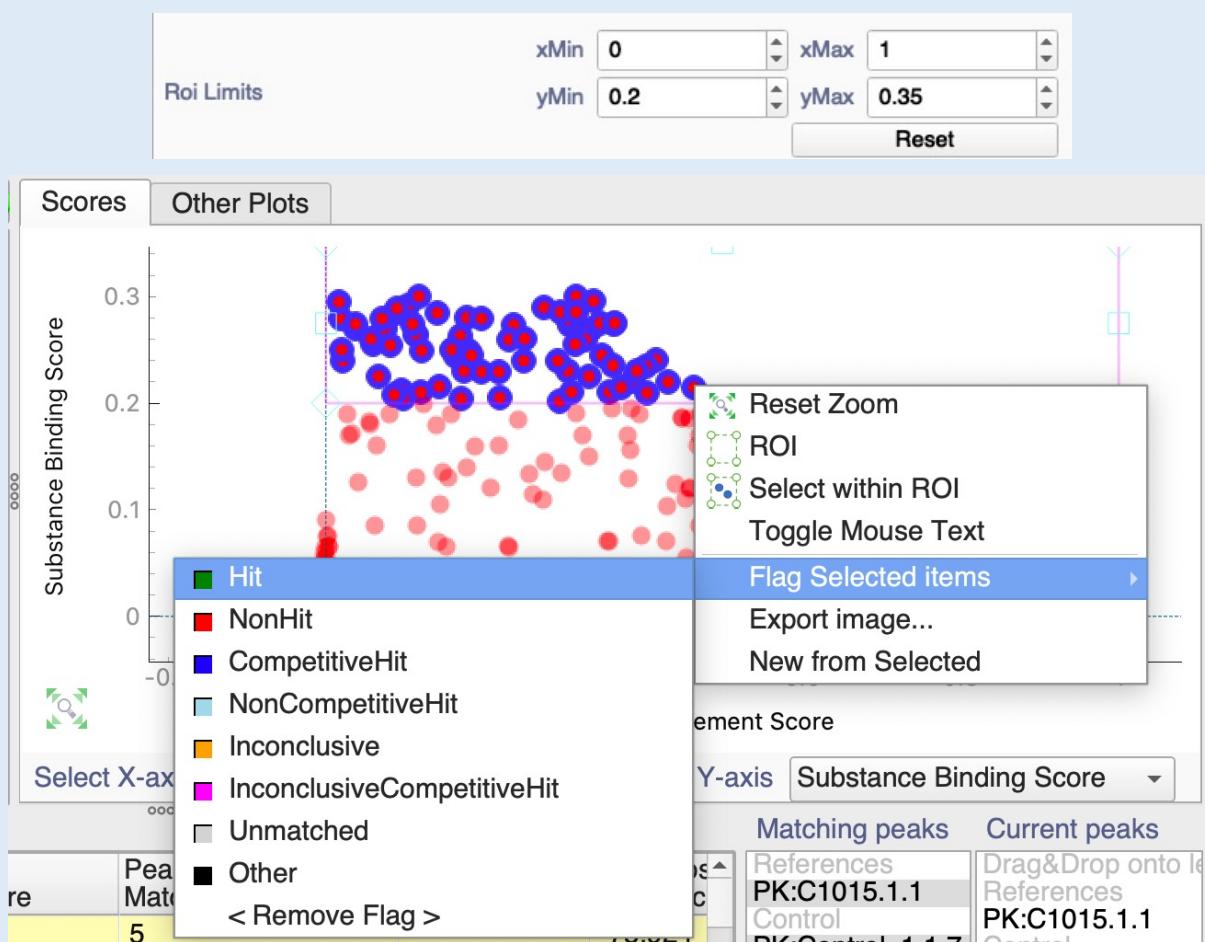


- Reset the view with

Additional options are available in the **Appearance** tab of the **Settings**:

Roi Limits	xMin: 0	xMax: 0
	yMin: 0	yMax: 0
<input checked="" type="checkbox"/> Link Roi with selection box		
Reset		
ROI line colour	<input type="color" value="#dimgrey"/>	
Scatter point colour	<input type="color" value="red"/>	
Scatter point size	10	
Scatter point Symbol	<input type="radio"/> circle	

Further types of plots including the molecular structure of the substances (if SMILES were entered into the project) are available in the **Other Plots** tab.



Before flagging Substances, always inspect the matches with the table selections. Snap peaks with the shortcut SE, or correct matches from the two list widgets at the bottom right corner.

5B Define Binding Hit Thresholds

- View mode: Substances
- Scatter Plot:
 - X-axis: Substance Displacement Score
 - Y-axis: Substance Binding Score

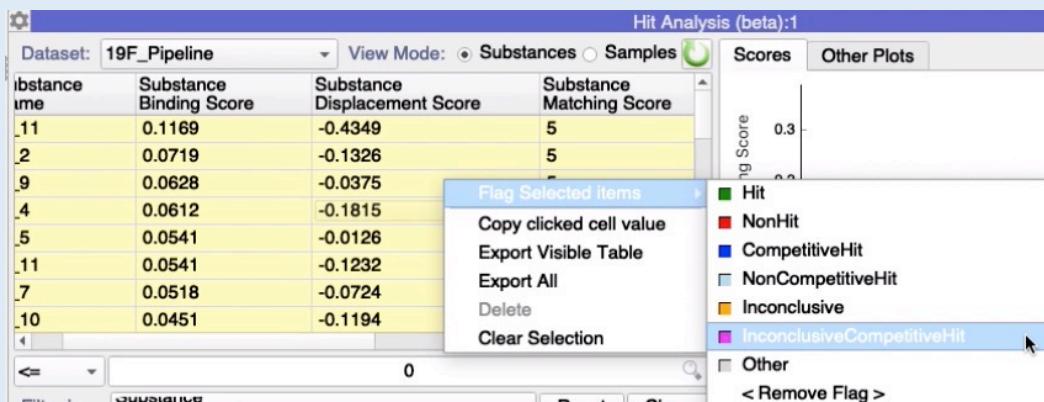
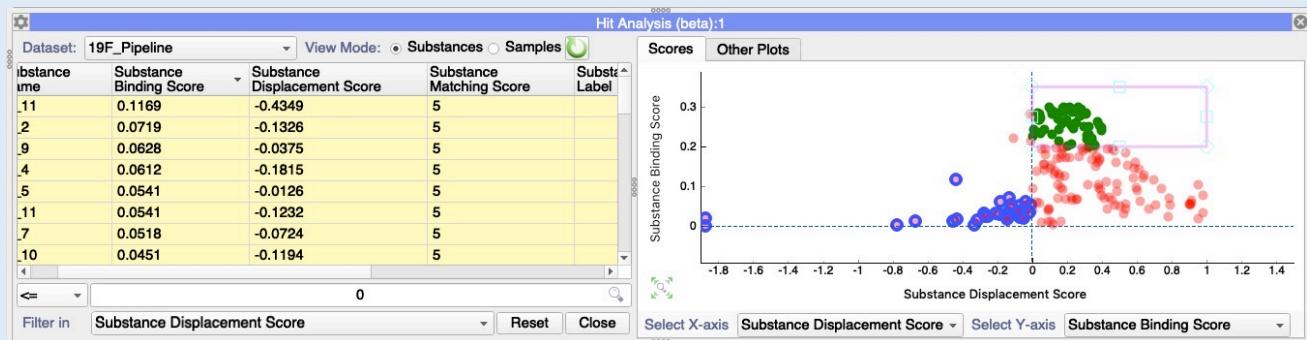
Open Settings:

- Calculation tab
 - Engine: AbsoluteRelativeChange
- Appearance tab
 - Roi Limits: click Reset
 - xMin: 0 xMax: 1
 - yMin: 0.2 , yMax: 0.35

This will create a **Region of Interest** on the scatter plot.

On the scatter plot:

- right-click on an item in the plot → Select within ROI
- right-click on an item in the plot → Flag Selected Items → Hit



The Hit analysis module has several filters for defining hits based on dynamically set threshold limits.

5c Filter Tables

Another way of defining hits is by applying one or multiple filters from the substance table:

- Right-click on the Substance Table header → Filter... or use shortcut FT

first filter:

- select: “<= “ 0.2 (less than)
- filter in: Substance Binding Score
- press the search button

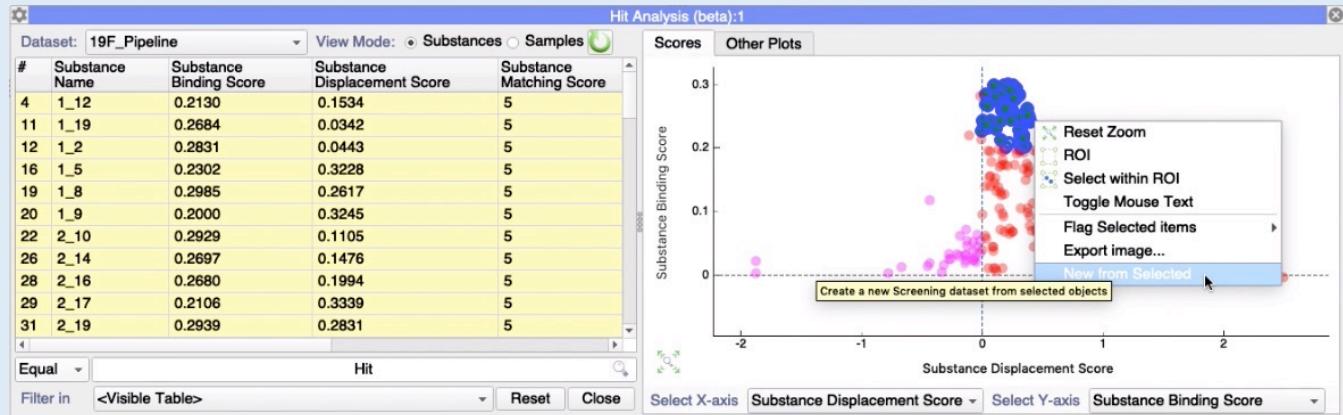
second filter:

- select: “<= “ 0
- filter in: Substance Displacement Score
- press the search button

- Select all rows

If a display is open, a warning will pop up: Click **No** to not add all the selected spectra on the current Spectrum Display.

- right-click on a row: **Flag Selected Items** → **Inconclusive Competitive Hit**



6A Extract and export

- Sort the substance table by **Substance Label** (click on the column header) or filter by **Equal “Hit”** in Visible Table or Substance Label (reset any previous filter first)
- Select all the rows for Substances flagged as **Hit**
- Move to the scatter plot, **right click** → **New from selected**
This will create a new dataset containing only this subset of substances. You can rename the dataset in the sidebar under **DataTables**.
- Select the newly created dataset on the Hit Analysis module, by selecting it from the **Dataset** dropdown menu.

Either continue to inspect the data as shown in Sections 4 and 5 or export the table:

- Right-click** on a Substance Table row → **Export All Columns**

6B Export raw data

To export the raw data which is used to build the **Hit Analysis Module**:

- Right-click** on a Substance Table row → **Export Raw Data**
- In the file dialog, Save as: type a name plus the extension, e.g.: .xlsx, .csv, .tsv or .json

This will create a table which contains all peak metadata, including peak ppm positions, heights etc. for further inspection or macros.

Create a NEF file

All the information regarding Samples, Substances, Spectra and Spectrum Groups can be contained within an NMR Exchange Format (NEF) file.

A NEF file can be created from an existing screening dataset with a few steps. These include **renaming** Samples, Spectra and Spectrum Groups from **Control** to **ReferenceMixtures** (or similar); **deleting** all other spectra and samples and finally **linking** the Substances to the Reference mixtures.

```

MacroEditor
[Icons: Open, Save, PDF, Run, Undo, Redo, Stop]
/Users/vad5/Documents/Tutorials/data/dataFeb2022/CcpnTutorialDataScreeningMarch22/ScreenTutorialMa
28
29 -> for obj in project.spectra:
30     name = obj.name
31     newName = name.replace('Control','RefMix')
32     obj.rename(newName)
33
34
35

```

7A Load project

- Continue with your Hit Analysis project or load the **dataset_2_completed.ccpn** project from the **ScreenTutorial/19F/dataset_2/** directory.

7B Renaming Control Samples, Spectra and Spectrum Group

- Double-click on the **SG:Controls** Spectrum Group in the sidebar and rename it to **RefMix**

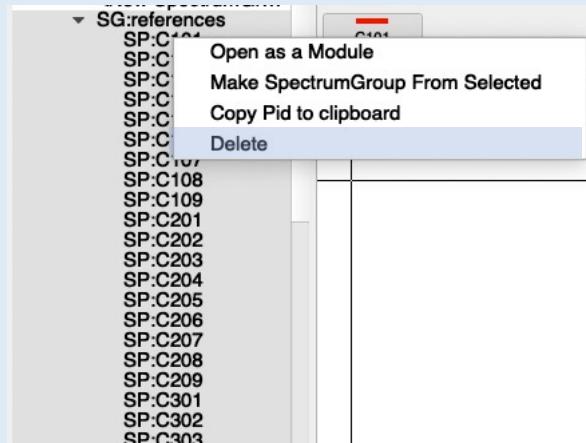
To rename the Control Spectra and Samples we will use a macro.

- Go to **Main Menu** → **Macro** → **New Macro Editor**
- Click on the **Open File** icon and select the **ScreenTutorial/19F/dataset_3/macros/RenameSpectra.py** file (or drag and drop the file onto the Macro Editor).
- Click on the **Play** button to run the macro. This may take a few moments.
- Now open the **ScreenTutorial/19F/dataset_3/macros/RenameSpectra.py** file and run this macro, too.

You should see in the sidebar that all Control Spectra and Samples have been renamed.

You can set your preferred Macro directory path in **File / Preferences / General**.

Create a NEF file



Python Console

```

SP:C001 , SP:C002 , SP:C003 , SP:C100_1 , SP:C101
project.deleteObjects('SP:Target_12', 'SP:Target_20', 'SP:Target_11', 'SP:Target_13', 'SP:Target_15', 'SP:Target_17'
project.deleteObjects('SP:Displacer_20', 'SP:Displacer_18', 'SP:Displacer_12', 'SP:Displacer_19', 'SP:Displacer_17', 'SP:Displacer_15', 'SG:Displacer')

In [6]: sg = get('SG:References')

In [7]: project.deleteObjects(*list(sg.spectra)+[sg])

In [8]: sg = get('SG:Target')

In [9]: project.deleteObjects(*list(sg.spectra)+[sg])

In [10]: sg = get('SG:Displacer')

In [11]: project.deleteObjects(*list(sg.spectra)+[sg])

In [12]:

```

7C Delete Reference, Target and Displacer Data

In the Sidebar:

- Expand the **SG:References** SpectrumGroup tree
- Select all spectra in the **SG:References** Spectrum Group, including the SpectrumGroup **SG:References** itself: **right-click → Delete**

Or

- If it isn't open already, open the **Python Console** by pressing the **Spacebar** twice.
- Run the commands:

```

sg = get('SG:References')
project.deleteObjects(*list(sg.spectra)+[sg])

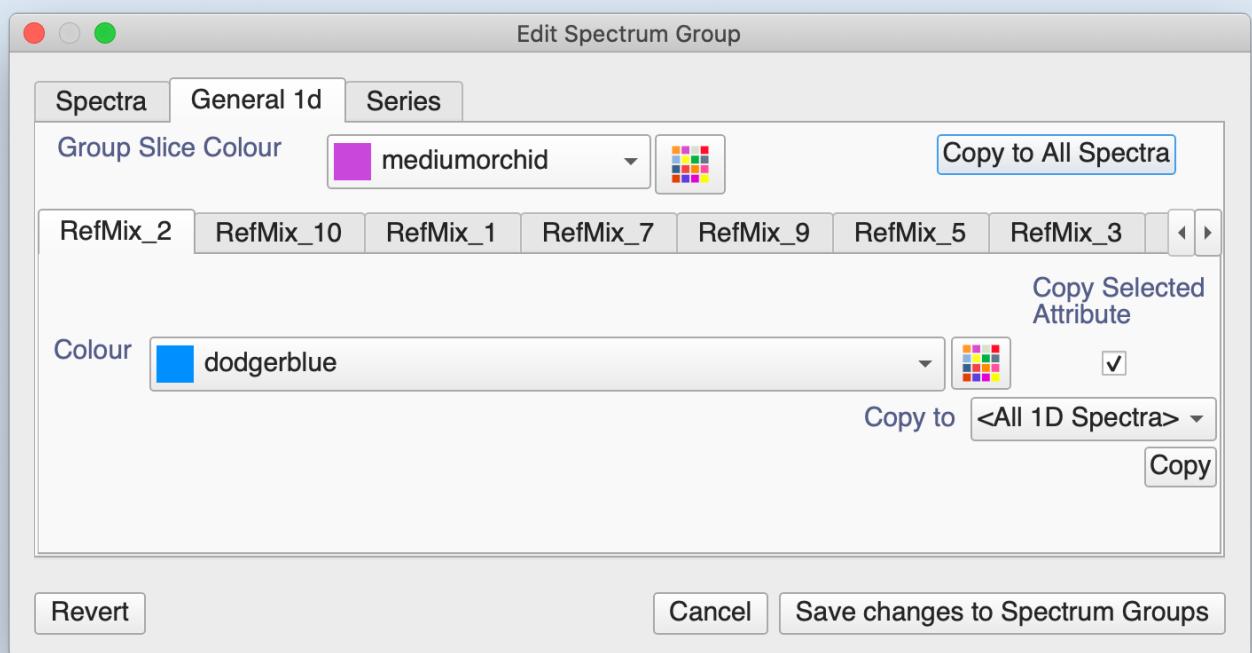
```

- Repeat for the **SG:Target** and **SG:Displacer** Spectrum Groups ammending the first list of code if you are using. it.
- Expand the **Samples** branch in the sidebar and select all **Target** and **Displacer** samples and delete with **right-click → Delete**

7D Link Substances to Specta

- Go to **Main Menu -> Macro -> New Macro Editor**
- Click on the **Open File** icon and select the **ScreenTutorial/19F/dataset_3/macros/LinkSubstancesToSpectra.py** file.
- Click on the **Play** button to run the macro.

Create a NEF file



7E Recolour Reference Mixture Spectra

- If you wish, change the colour of the spectra in the new **SG:RefMix** Spectrum Group as shown in **Section 2E** by double-clicking on **SG:RefMix** in the sidebar, going to the **General 1d** tab, selecting a new **Group Slice Colour** and pressing **Copy to All Spectra** followed by either **Cancel** or **Save changes to Spectrum Groups**.

7F Correct peaks (optional for this tutorial)

Because the NEF file will function as a template for future screening analyses, it is wise to inspect all Reference Mixtures, ensuring all peaks are correctly annotated with the respective Reference Substance:

- In the sidebar, go to the first sample **SA:RefMix_1** and **right-click → Open Linked spectra** or drag & drop it into the Drop Area.
- Make sure the **RefMix_1** peaks are correctly positioned compared to the references, or use **SE** to re-snap the selected peak(s) to their extremum. (You can change snapping limits in Preferences, Spectrum Tab, **1d Search Box Widths**)
- In the sidebar, use the shortcut **Ctrl (or Cmd)+up/down** directional keys to visualise the next/previous sample and associated spectra.

Peak annotation

When using NEF files as a template for screening calculations, Substances are tracked in the Reference Mixture signals through the **Peak Annotations**.

Peak Annotation names can be made in three parts:

Prefix	substance name	The exact substance name	Mandatory
Separator	_	Underscore	Mandatory if Suffix
Suffix	Any	Any single word tag that can help identify the signal; e.g.: impurity, TFA, Salt, a serial number, an atom name etc.	Optional

Peak annotation examples: **Compound1_CF3**, **Compound2_Salt**, **Unknown**.

Peaks with a figure of merit of 0 are excluded from screening calculations.

7G Add extra annotations for impurities, solvents etc and exclude them from calculations

- Close all GUI Modules.
- Drag and drop the **SG:RefMix** Spectrum Group from the Sidebar into the Drop Area.
- Pick and select all peaks in the region -99.9, -100.1 ppm
- Open the **Macro Editor** from **Macro → New Macro Editor**
- Open the **ScreenTutorial/19F/dataset_3/macros/AnnotateCurrentPeaks.py** file or drag and drop it onto the Macro Editor.
- Run with the play button (while the peaks are selected).

A code snippet is shown below

```
tag = 'Impurity'

for peak in current.peaks:
    peak.annotation = '_'.join(filter(None, set([peak.annotation, tag])))
    peak.figureOfMerit = 0
```

Warning Copy&Paste code from PDF might lose the original indentation causing syntax errors

Create a NEF file



7H Change peak annotation colours

You can change the peak symbol/text colours so that real signal peaks (black) are graphically distinguishable from excluded (light grey) ones:

- Open the **ScreenTutorial/19F/dataset_3/macros/ColourPeakLists.py** macro in the Macro Editor and run.

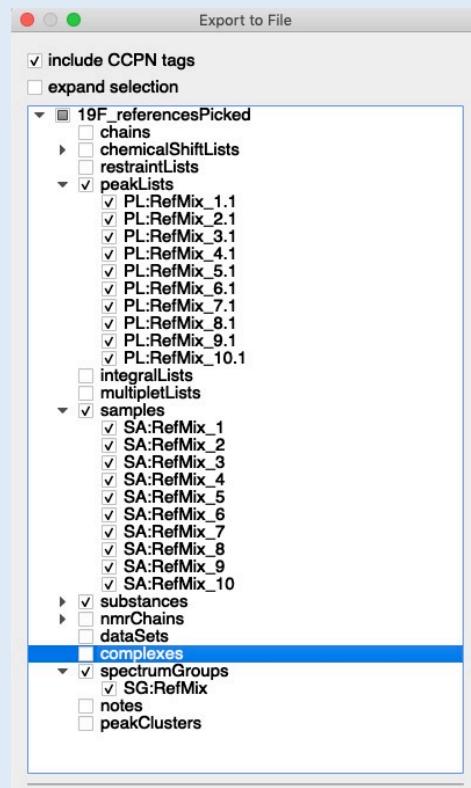
A code snippet is shown below

```
black = '#000000'
lightGrey = '#b7b7b7'

for peakList in project.peakLists:
    peakList.symbolColour = black
    peakList.textColour = black
    peakList.meritThreshold = 0.5
    peakList.meritEnabled = True
    peakList.meritColour = lightGrey
```

Warning Copy&Paste code from PDF might lose the original indentation causing syntax errors

Create a NEF file



7J Export to NEF

Export metadata related to the Reference Mixtures:

- **Main Menu → Project → Export → NEF File** (or use shortcut EX)
- In the NEF Dialog untick all first, and tick only the following:

General:

- include CCPN Tags,

Project tree:

- PeakLists
- Samples
- Substances
- SpectrumGroups

- Save to your local disk



8A Import From NEF

The NEF file contains information about your reference mixtures, including peak lists, samples and substances as well as some formatting.

- Select the **19F_MixtureReferences.nef** located in the **ScreenTutorial/19F/dataset_3/** directory and drag it onto the sidebar.
- When prompted, select to open as **New Project**.
- Change the **SG:RefMix** spectrum group spectrum colours as shown in **Section 2E** if desired.

Import Data from Excel and NEF

	A	B	C	D	E	F	G	H	I	J	K
1	sampleName	group	spectrumPath	comment	pH	ionicStrength	amount	amountUnit	hazardou		
2	0	Control_1	Control	Control_1 Control_1	8	8.5	9.133	g			
3	1	Control_2	Control	Control_2 Control_2	6	8.5	7.571	g			
4	2	Control_3	Control	Control_3 Control_3	7	7	8.078	g			
5	3	Control_4	Control	Control_4 Control_4	7	7.5	8.878	g			
6	4	Control_5	Control	Control_5 Control_5	9	8.5	9.731	g			
7	5	Control_6	Control	Control_6 Control_6	6	7	8.176	g			
8	6	Control_7	Control	Control_7 Control_7	7	7.5	9.475	g			
9	7	Control_8	Control	Control_8 Control_8	7	9	9.955	g			
10	8	Control_9	Control	Control_9 Control_9	7.5	7	7.603	g			
11	9	Control_10	Control	Control_10 Control_10	9	8	7.964	g			
12											

SampleControl SampleTarget SampleDisplacer +

Ready 180%

8B Import Data from Excel

Import the latest screening data from the Excel file

- Load the Excel file from the **dataset_3** tutorial data folder:

ScreenTutorial/19F/dataset_3/Data_Time_x/lookup_19F_TimeX.xlsx

This lookup only contains the **Sample** sheets for the Control, Target and Displacer data without the **SampleComponents** field and the **Substances** sheet.

This information has already been imported into the project from the NEF file.



Automated Peak Matching from Reference Mixtures



9A Load Data

Import your data as shown in **Section 8**.

9B Setup screening pipeline

- Open the Pipeline module (**Menu → Screen → Pipeline** or shortcut PI)
- In the sidebar, multiselect SpectrumGroups **SG:Control**, **SG:Target**, **SG:RefMix**
- drag and drop them into the pipeline **Input Data** box
- In the list of **Pipes** search for and add to the pipeline area:

1. Noise Threshold

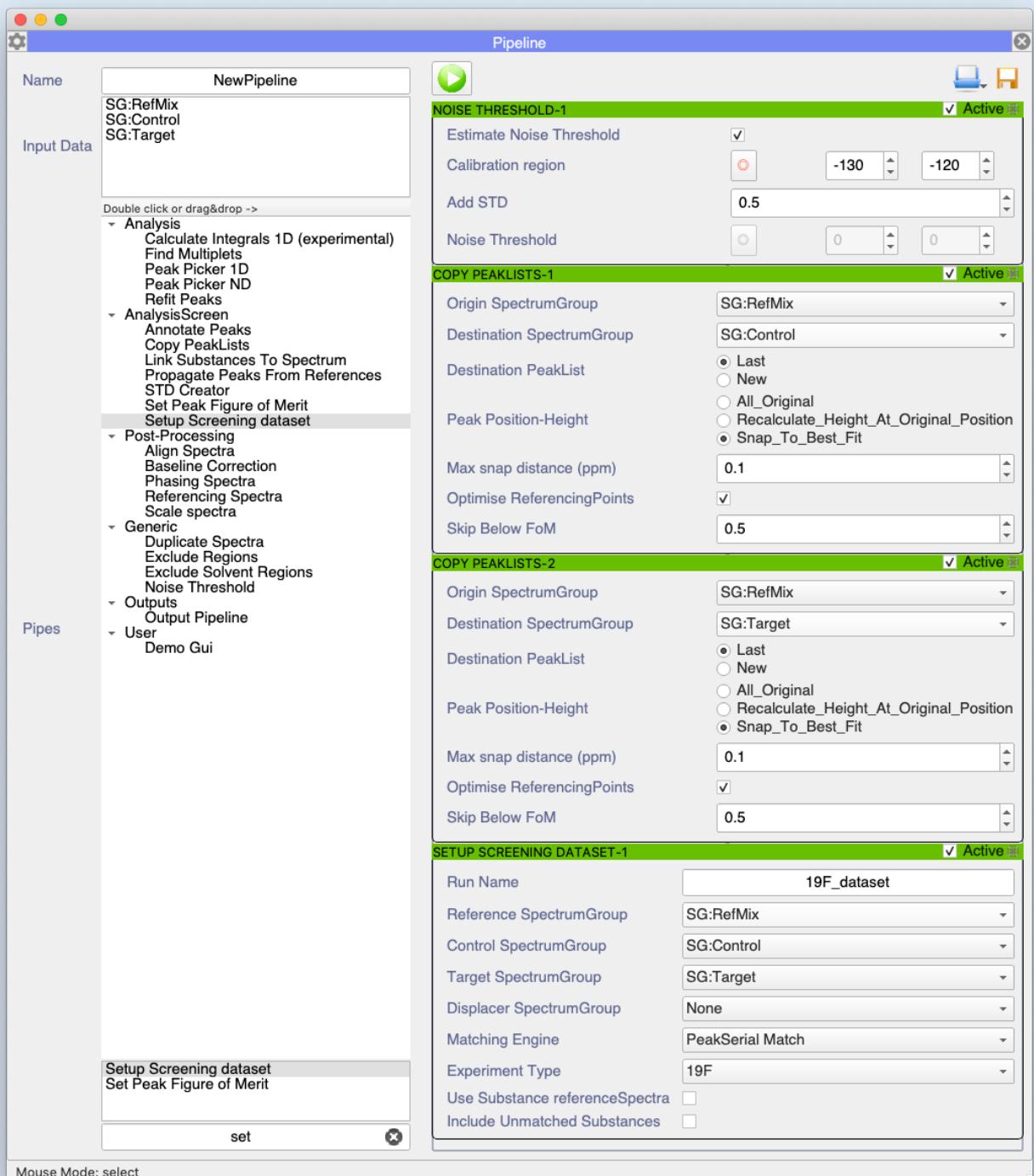
- tick **Estimate Noise Threshold**
- Calibration region -130, -120

2. Copy PeakLists (1)

- Origin SpectrumGroup: **SG:RefMix**
- Destination SpectrumGroup: **SG:Control**

... continued

Automated Peak Matching from Reference Mixtures



...continued

add the following pipes

3. Copy PeakLists (2)

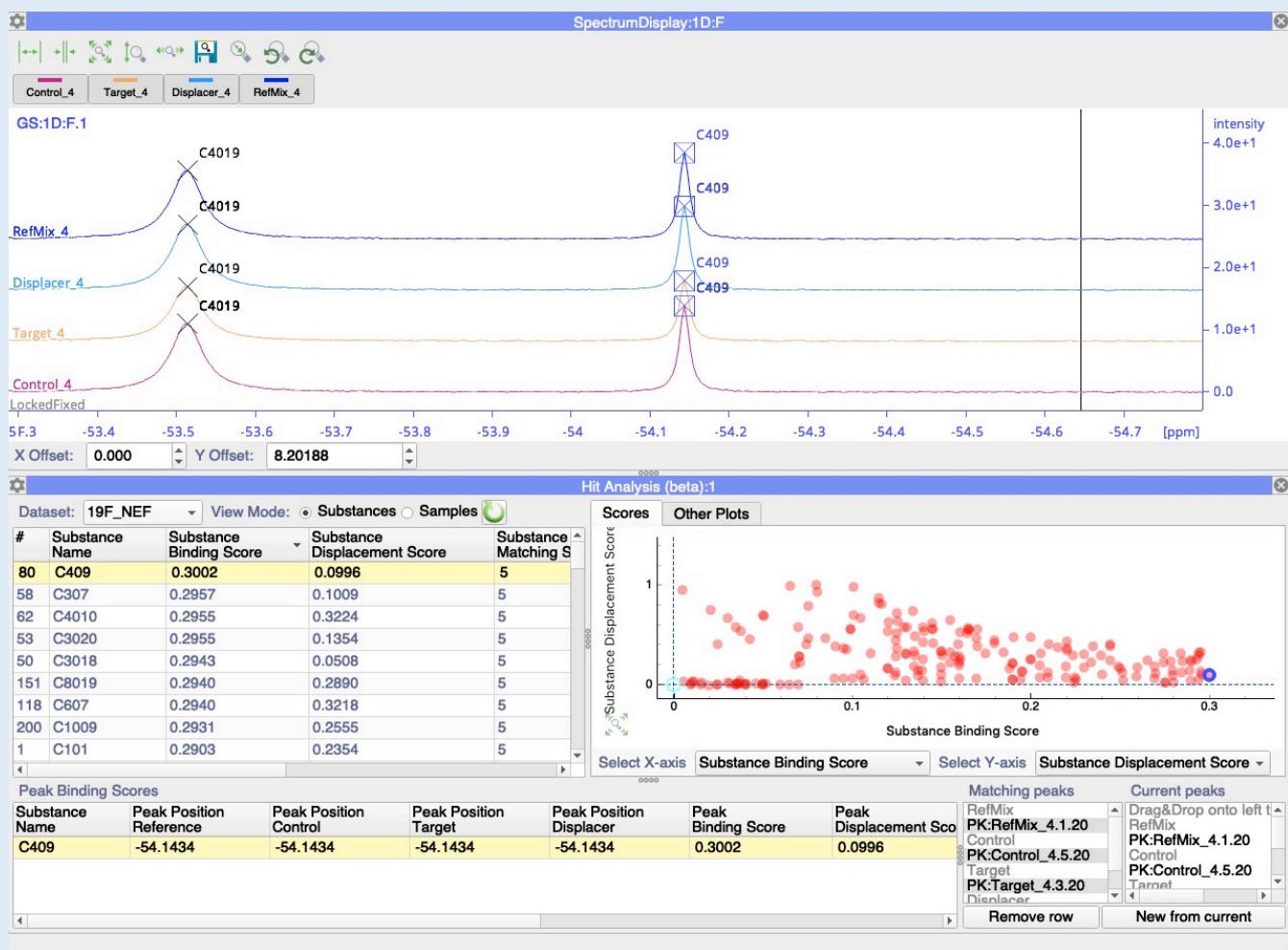
- Origin SpectrumGroup: SG:RefMix
- Destination SpectrumGroup: SG:Target

4. Setup ScreenDataset

- **Run name:** 19F_dataset
- **Reference SpectrumGroup:** SG:RefMix
- **Select SG:Control and SG:Target** for their respective entries
- **Matching engine:** Peak Serial Match
- **untick Use_Substance_ReferenceSpectra**

- Run the pipeline.
- Inspect the results in the Hit Analysis Module as shown in Sections 4–6.

Automated Peak Matching from Reference Mixtures



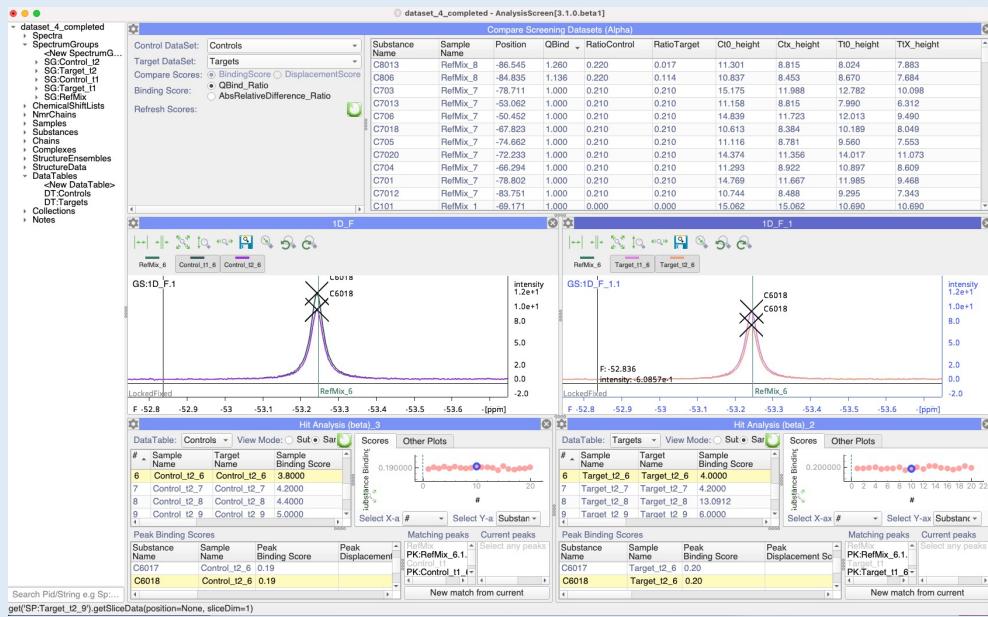
9C Completed Project

An example of a completed project is available in the dataset_3 directory:

.../19F/dataset_3/19F_NEF_Completed.ccpn

Hit Analysis Module

When selecting Substance or Sample items in the main tables, you may notice how the Singleton Reference Spectra are replaced by the Reference Mixture spectra. The peak annotations provide a visual reference to the matching Substance.



10A Open data in program

In this section you will compare two screens performed on the same library at different CPMG times.

You can now either recreate the project in **Sections 10B-D** or load a project which has already two completed screening datasets and skip to **Section 10E**:

- Drag and drop the **ScreenTutorial/19F/dataset_4/dataset_4_completed.ccpn** folder in the sidebar or Drop Area

10B Setup two parallel screens (optional)

From the **ScreenTutorial/19F/dataset_4/** folder:

- Load the NEF file, as a new project.
- Load the Excel files in the directory Data_Time_1 and Data_Time_2.

10C Copy peakLists (optional)

- Open the pipeline module (**PI**), add as input data all Spectrum Groups.

Use the pipe **Copy Peak Lists** pipe:

- Origin SpectrumGroup: SG:RefMix
- Destination SpectrumGroup: SG:Control_t1
- Keep all settings as default.

Repeat this pipe for the other spectrumGroups:

- SG:RefMix → SG:Target_t1
- SG:RefMix → SG:Control_t2
- SG:RefMix → SG: Target _t2 (see image next page)

Run the pipeline, once completed close all pipes and keep the pipeline opened.

10 Using data with multiple CPMG delays

COPY PEAKLISTS-1 ✓ Active

Origin SpectrumGroup	SG:RefMix
Destination SpectrumGroup	SG:Control_t1
Destination PeakList	<input checked="" type="radio"/> Last <input type="radio"/> New <input type="radio"/> All_Original <input type="radio"/> Recalculate_Height_At_Original_Position <input checked="" type="radio"/> Snap_To_Best_Fit
Peak Position-Height	0.1
Max snap distance (ppm)	0.1
Optimise ReferencingPoints	<input checked="" type="checkbox"/>
Skip Below FoM	0.5

COPY PEAKLISTS-2 ✓ Active

Origin SpectrumGroup	SG:RefMix
Destination SpectrumGroup	SG:Target_t1
Destination PeakList	<input checked="" type="radio"/> Last <input type="radio"/> New <input type="radio"/> All_Original <input type="radio"/> Recalculate_Height_At_Original_Position <input checked="" type="radio"/> Snap_To_Best_Fit
Peak Position-Height	0.1
Max snap distance (ppm)	0.1
Optimise ReferencingPoints	<input checked="" type="checkbox"/>
Skip Below FoM	0.5

COPY PEAKLISTS-3 ✓ Active

Origin SpectrumGroup	SG:RefMix
Destination SpectrumGroup	SG:Control_t2
Destination PeakList	<input checked="" type="radio"/> Last <input type="radio"/> New <input type="radio"/> All_Original <input type="radio"/> Recalculate_Height_At_Original_Position <input checked="" type="radio"/> Snap_To_Best_Fit
Peak Position-Height	0.1
Max snap distance (ppm)	0.1
Optimise ReferencingPoints	<input checked="" type="checkbox"/>
Skip Below FoM	0.5

COPY PEAKLISTS-4 ✓ Active

Origin SpectrumGroup	SG:RefMix
Destination SpectrumGroup	SG:Target_t2
Destination PeakList	<input checked="" type="radio"/> Last <input type="radio"/> New <input type="radio"/> All_Original <input type="radio"/> Recalculate_Height_At_Original_Position <input checked="" type="radio"/> Snap_To_Best_Fit
Peak Position-Height	0.1
Max snap distance (ppm)	0.1
Optimise ReferencingPoints	<input checked="" type="checkbox"/>
Skip Below FoM	0.5

SETUP SCREENING DATASET-1 ✓ Active

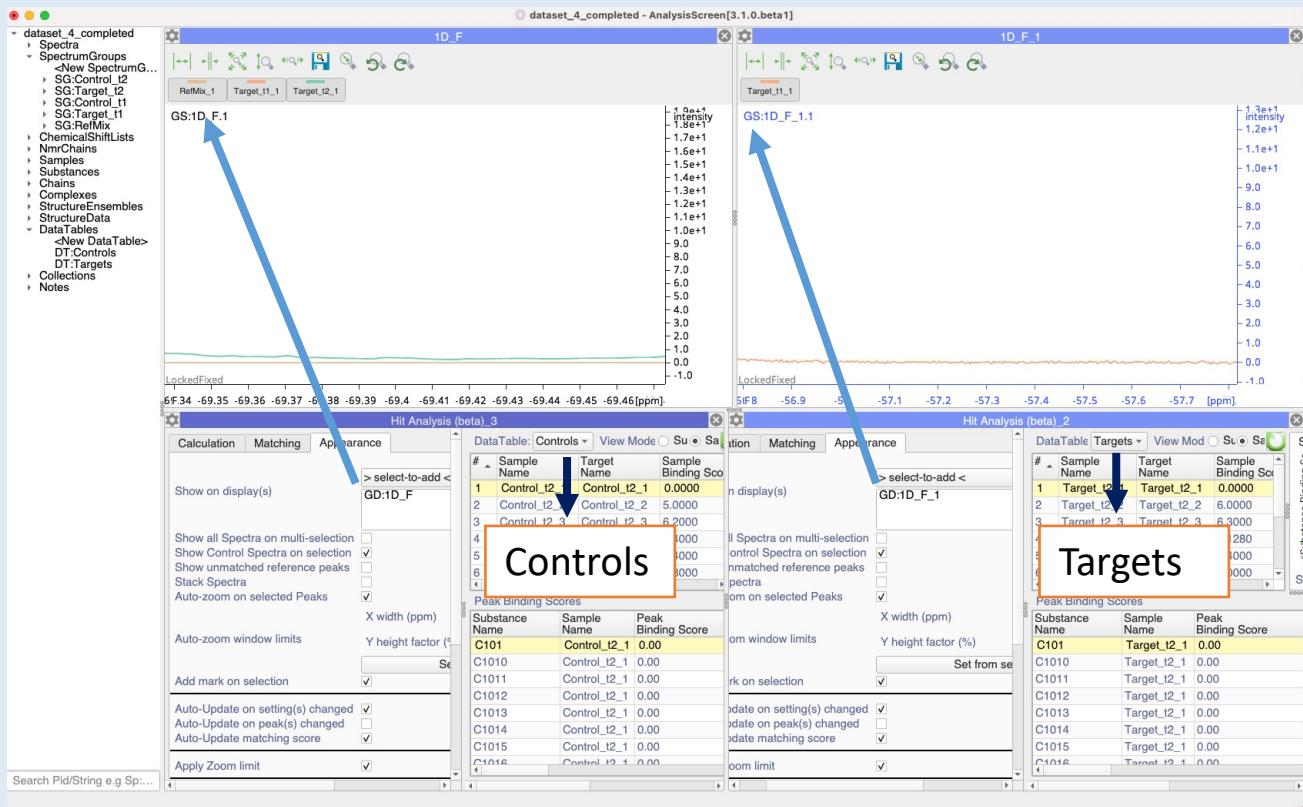
Run Name	Controls
Reference SpectrumGroup	SG:RefMix
Control SpectrumGroup	SG:Control_t1
Target SpectrumGroup	SG:Control_t2
Displacer SpectrumGroup	None
Matching Engine	PeakSerial Match
Experiment Type	19F
Use Substance referenceSpectra	<input type="checkbox"/>
Include Unmatched Substances	<input type="checkbox"/>

SETUP SCREENING DATASET-2 ✓ Active

Run Name	Targets
Reference SpectrumGroup	SG:RefMix
Control SpectrumGroup	SG:Target_t1
Target SpectrumGroup	SG:Target_t2
Displacer SpectrumGroup	None
Matching Engine	PeakSerial Match
Experiment Type	19F
Use Substance referenceSpectra	<input type="checkbox"/>
Include Unmatched Substances	<input type="checkbox"/>

10D Setup Screening Datasets (optional)

- Add the **Set up Screening dataset** pipe twice.
- Set the first pipe up for the Control data (as shown above):
 - Run Name: **Controls**
 - Reference SpectrumGroup: **SG:RefMix**
 - Control SpectrumGroup: **SG:Control_t1**
 - Target SpectrumGroup: **SG: Control_t2** (not the target!)
 - Displacer SpectrumGroup: **None**
 - Matching Engine: **PeakSerial Match**
 - Experiment type: **19F**
 - Untick the last two checkboxes.
- Repeat for the Target data with **Targets**, **SG:Target _t1** and **SG:Target _t2** for the Run Name, Control SpectrumGroup and Target SpectrumGroup, respectively.
- Run the pipeline and then close the module.



10E

Set up displays and HitAnalysis Modules

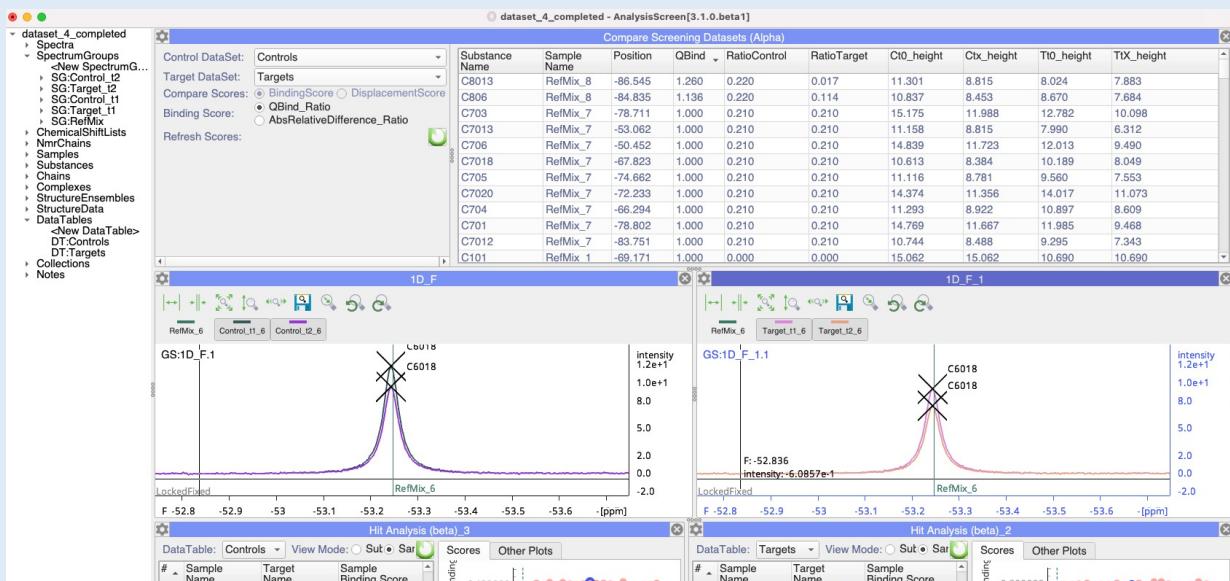
- Open two SpectrumDisplays and two HitAnalysis Modules as shown above.

In the Hit Analysis modules:

- Set the **DataTable** to **Controls** on the left and to **Targets** on the right.
- Open the **Settings** panels: Set **Show on display(s)** to **GD:1D_F** on the left and to **GD:1D_F_1** on the right.

You will not need to apply any other settings in the Hit Analysis modules. You might like to reduce their size to a minimum height.

10 Using data with multiple CPMG delays



Q^{bind} score

One way of comparing multiple CPMG times is by using the Q^{bind} score*

$$Q^{\text{bind}} = \frac{\text{Intensity Ratio} + ^{\text{Target}}}{\text{Intensity Ratio} ^{\text{Control}}}$$

with:

$$\text{Intensity Ratio} = \frac{\text{Peak Intensity}^{\text{time } x}}{\text{Peak Intensity}^{\text{time } 0}}$$

A lower value such as $Q^{\text{bind}} < 0.32$ may reflect a strong binding event; values between 0.33–0.66 a medium binding, whereas $Q^{\text{bind}} > 0.67$ may indicate no interaction between the target and the molecule.

* 19F NMR-Based Fragment Screening for 14 Different Biologically Active RNAs and 10 DNA and Protein Counter-Screens. Binas et al. *ChemBioChem* 2021, 22, 423 – 433. doi.org/10.1002/cbic.202000476

10F Open Compare Screening Datasets

If not already present, open the module **Compare Screens** from the main menu:

- Main Menu → Screen → Compare Screens.
- Open the Settings panel, select the **Control** and **Target** datasets and select **Q^{bind} Ratio** as the **Binding Score**.
- Click the refresh button to populate the table.
- On the table, filter by the **QBind** score to identify potential strong binders:
- Right click the header and click **Filter...** (or use shortcut **FT**).

Filter between 0.01 and 0.32 in in **Qbind**.

- Click **Search**.

The table will now display only a few entries. Inspect the spectra by selecting the row. The displays will navigate accordingly to the peaks.

dataset_4_completed

- Spectra
- SpectrumGroups
 - <New SpectrumG...
 - SG:Control_I2
 - SG:Target_I2
 - SG:Control_I1
 - SG:Target_I1
 - SG:RefMix
- ChemicalShiftLists
- Notes
- Samples
- Substances
- Chem
- Complexes
- StructureEnsembles
- StructureData
- DataTables
 - <New DataTable>
 - DT:Controls
 - DT:Targets
- Collections
- Notes

dataset_4_completed - AnalysisScreen[3.1.0.beta1]

Compare Screening Datasets (Alpha)

Control DataSet:	Targets	Substance	Sample Name	Position	QBind	RatioControl	RatioTarget	Ct0_height	Ctx_height	Tt0_height	Ttx_height
Controls	Targets	C1001	RefMix_10	-52.433	0.127	0.180	0.896	13.043	10.696	11.020	1.146
		C1005	RefMix_10	-63.684	0.159	0.180	0.870	13.234	10.852	11.382	1.480
		C468	RefMix_4	-79.521	0.162	0.120	0.840	12.561	11.054	10.236	1.638
		C4011	RefMix_4	-72.772	0.218	0.120	0.808	12.998	11.438	11.242	2.159

Refresh Scores: Between 0.01 0.32 Filter in QBind Search Reset Close

1D_F 1D_F_1

GS:1D_F_1 GS:1D_F_1.1

Hit Analysis (beta).3 Hit Analysis (beta).2

DataTable: Controls View Mode Sul Sa Scores Other Plots # Sample Name Target Name Sample Binding Score Inset Bin 0 180000 0 10 20 7 Control_I2_7 Control_I2_7 4.2000 #

DataTable: Targets View Mode Sul Sa Scores Other Plots # Sample Name Target Name Sample Binding Score Inset Bin 0 0 2 4 6 8 10 12 14 16 18 20 22 7 Target_I2_7 Target_I2_7 4.2000 #

Search Pid/String e.g Sp... get(SP:Target_I2_9).getSliceData(position=None, sliceDim=1)

Strong Binders
QBind < 0.32

dataset_4_completed

- Spectra
- SpectrumGroups
 - <New SpectrumG...
 - SG:Control_I2
 - SG:Target_I2
 - SG:Control_I1
 - SG:Target_I1
 - SG:RefMix
- ChemicalShiftLists
- Notes
- Samples
- Substances
- Chem
- Complexes
- StructureEnsembles
- StructureData
- DataTables
 - <New DataTable>
 - DT:Controls
 - DT:Targets
- Collections
- Notes

dataset_4_completed - AnalysisScreen[3.1.0.beta1]

Compare Screening Datasets (Alpha)

Control DataSet:	Targets	Substance	Sample Name	Position	QBind	RatioControl	RatioTarget	Ct0_height	Ctx_height	Tt0_height	Ttx_height
Controls	Targets	C8020	RefMix_8	-66.832	0.359	0.220	0.720	14.323	11.172	10.602	2.969
		C8014	RefMix_8	-61.883	0.359	0.220	0.720	13.711	10.694	11.313	3.168
		C8016	RefMix_8	-70.704	0.359	0.220	0.720	11.966	9.333	10.297	2.883
		C8012	RefMix_8	-54.501	0.359	0.220	0.720	14.319	11.169	13.813	3.868
		C8018	RefMix_8	-82.131	0.359	0.220	0.720	10.567	8.242	9.518	2.665
		C8011	RefMix_8	-63.951	0.359	0.220	0.720	13.754	10.728	11.142	3.120
		C803	RefMix_8	-51.262	0.359	0.220	0.720	11.006	8.584	8.038	2.251
		C808	RefMix_8	-50.543	0.359	0.220	0.720	10.688	8.337	9.364	2.622

Refresh Scores: Between 0.66 0.33 Filter in QBind Search Reset Close

1D_F 1D_F_1

GS:1D_F_1 GS:1D_F_1.1

Hit Analysis (beta).3 Hit Analysis (beta).2

DataTable: Controls View Mode Sul Sa Scores Other Plots # Sample Name Target Name Sample Binding Score Inset Bin 0 220000 0 10 20 8 Control_I2_8 Control_I2_8 4.4000 #

DataTable: Targets View Mode Sul Sa Scores Other Plots # Sample Name Target Name Sample Binding Score Inset Bin 0 0 2 4 6 8 10 12 14 16 18 20 22 8 Target_I2_8 Target_I2_8 13.0912 #

Search Pid/String e.g Sp... get(SP:Target_I2_9).getSliceData(position=None, sliceDim=1)

Medium Binders
QBind 0.33-0.66

dataset_4_completed

- Spectra
- SpectrumGroups
 - <New SpectrumG...
 - SG:Control_I2
 - SG:Target_I2
 - SG:Control_I1
 - SG:Target_I1
 - SG:RefMix
- ChemicalShiftLists
- Notes
- Samples
- Substances
- Chem
- Complexes
- StructureEnsembles
- StructureData
- DataTables
 - <New DataTable>
 - DT:Controls
 - DT:Targets
- Collections
- Notes

dataset_4_completed - AnalysisScreen[3.1.0.beta1]

Compare Screening Datasets (Alpha)

Control DataSet:	Targets	Substance	Sample Name	Position	QBind	RatioControl	RatioTarget	Ct0_height	Ctx_height	Tt0_height	Ttx_height
Controls	Targets	C6018	RefMix_6	-53.243	0.988	0.190	0.200	11.884	9.626	9.389	7.511
		C6017	RefMix_6	-60.354	0.988	0.190	0.200	12.566	10.178	11.680	9.344
		C6004	RefMix_6	-51.982	0.988	0.200	0.200	12.762	10.337	12.497	9.998
		C6014	RefMix_6	-67.552	0.988	0.190	0.200	13.267	10.746	10.612	8.489
		C6002	RefMix_6	-78.082	0.988	0.190	0.200	12.632	10.232	10.931	8.745
		C6010	RefMix_6	-51.714	0.988	0.190	0.200	11.509	9.322	10.025	8.020
		C6012	RefMix_6	-72.233	0.988	0.190	0.200	13.125	10.632	10.105	8.084
		C6013	RefMix_6	-65.932	0.988	0.190	0.200	11.416	9.247	9.130	7.304
		C6019	RefMix_6	-69.512	0.988	0.190	0.200	14.083	11.715	12.279	9.082

Refresh Scores: Between 0.67 0.99 Filter in QBind Search Reset Close

1D_F 1D_F_1

GS:1D_F_1 GS:1D_F_1.1

Hit Analysis (beta).3 Hit Analysis (beta).2

DataTable: Controls View Mode Sul Sa Scores Other Plots # Sample Name Target Name Sample Binding Score Inset Bin 0 190000 0 10 20 6 Control_I2_6 Control_I2_6 3.8000 #

DataTable: Targets View Mode Sul Sa Scores Other Plots # Sample Name Target Name Sample Binding Score Inset Bin 0 0 2 4 6 8 10 12 14 16 18 20 22 6 Target_I2_6 Target_I2_6 4.0000 #

Search Pid/String e.g Sp... get(SP:Target_I2_9).getSliceData(position=None, sliceDim=1)

Non Binders
QBind > 0.67

Contact Us

Website:

www ccpn.ac.uk

Suggestions and comments:

support@ccpn.ac.uk

Issues and bug reports:

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

Cite Us

Mureddu, L. et al. CcpNmr AnalysisScreen, a new software programme with dedicated automated analysis tools for fragment-based drug discovery by NMR. *J. Biomol. NMR* (2020)

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