MScreen Manual

Last Revision: 30-Sep-2014

MScreen 1.2

---

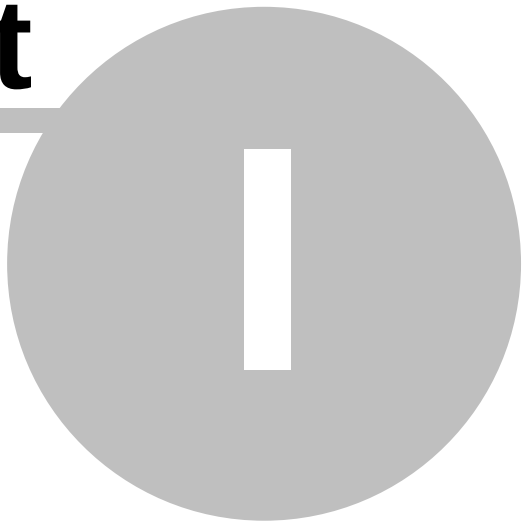
by MESTRELAB RESEARCH

This is the manual of MScreen 1.2

MScreen Manual

Table of Contents

Part



1 Intro

The automatic analysis of ligand screening experiments requires several conditions to be met, not only the reprocessing must be performed in optimal conditions of phase and baseline correction, but the data ought to be correctly referenced and aligned.

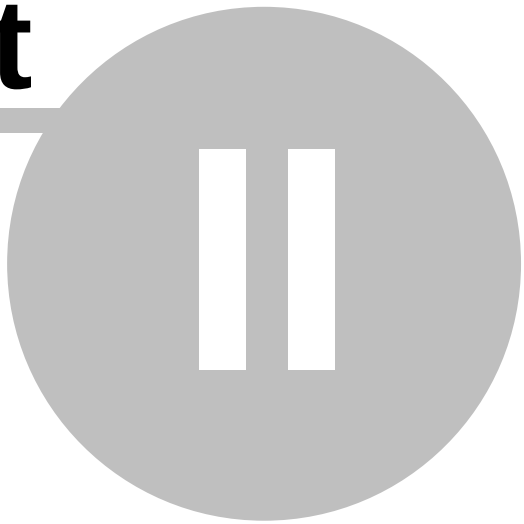
MScreen is a state of automatic analysis tool for ligand screening NMR data. It offers flexibility for the analysis and reporting of results.

The results viewer enables quick and efficient inspections

Amongst the main characteristics that we can found are:

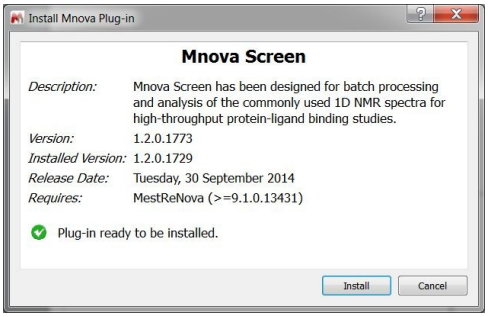
- Pattern recognition algorithms map your experimental data
- For 1H and 19F data
- Contemplates the use of STD, wLOGSY, T1ρ, CMPG experiment types
- Uses deconvolution to effectively pick peak even weak signals
- Automatically align and normalise between datasets to ensure effective analysis
- Result viewer to quickly inspect your hits
- Export metrics of your analysis into your favourite statistical package

Part

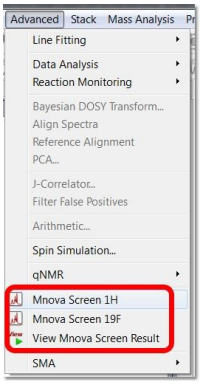


2 Installation

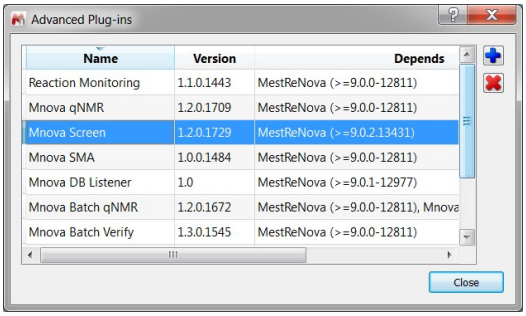
MScreen is a separate plugin and it comes as a separate installer, aside from the main Mnova installer, in order to proceed with the installation, you will need to download the corresponding Mnova-Screen-.zip installer, once this is done just drag and drop the installer over Mnova and you will be presented with the installation screen, click “Install” and the plugin will be installed.



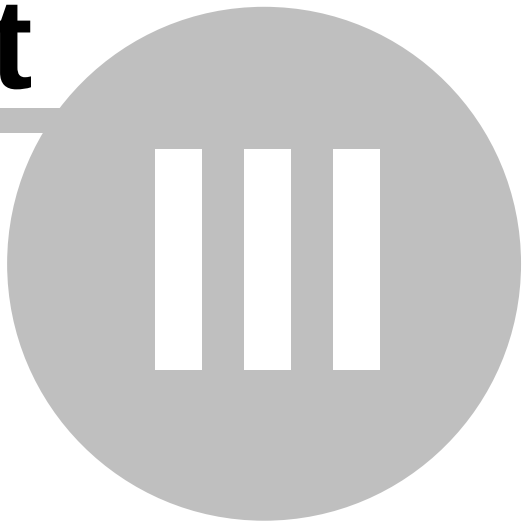
Once installed, MScreen will be found under the “Advanced Menu” of Mnova:



If at any time you need to check the version installed, just follow the menu 'Edit/Advanced Plugins' dialog:

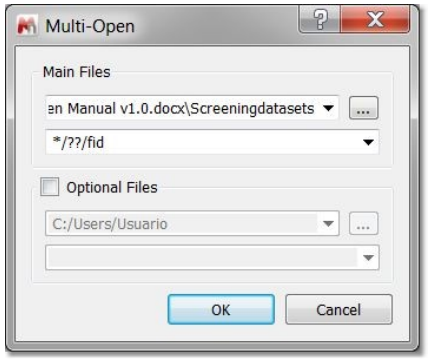


Part

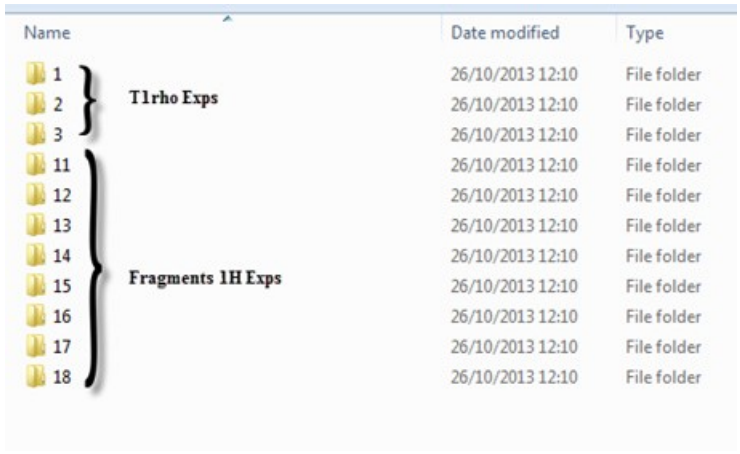


3 Preparation of reprocessing templates for each type of experiment.

In order to prepare the data to identify for the optimum reprocessing parameter, first we need to open each individual experiment, to do so effectively we will use the “Multiopen” script, this can be found at the menu “Scripts/Import/Multi-open”, we should see something like:



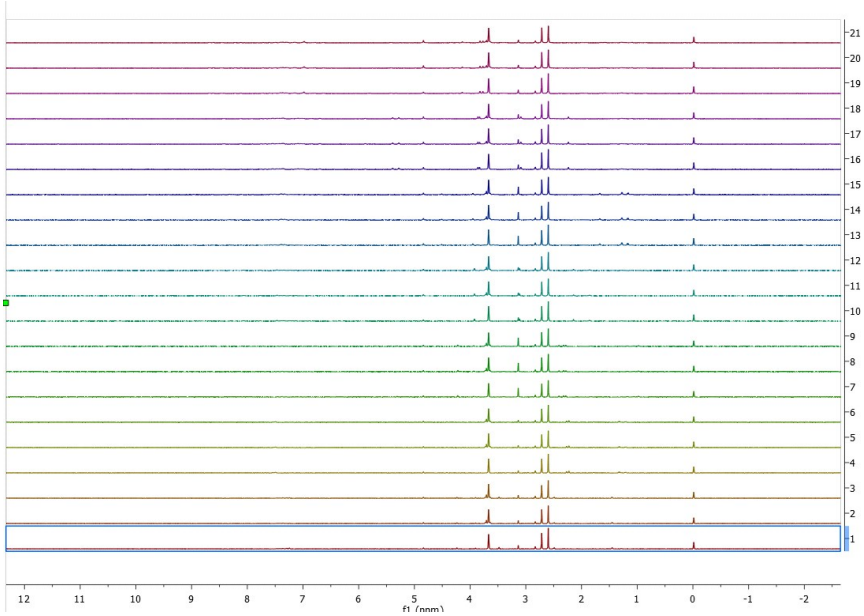
Next we need to point the script in the right direction so it will open only those spectra we desire, note that all the directories have a similar same structure:



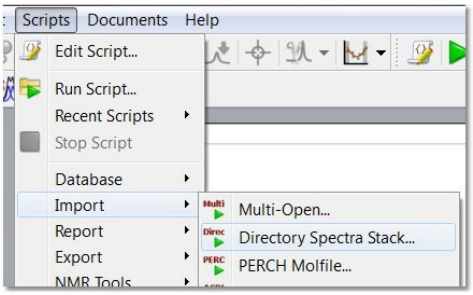
We can see that all of the T1rho experiments are on folders 1-3 and the individual fragments on folders 11 and up. Hence in order to map the data and be able to open both types of experiments separately we will have to use the following file masks:

- For T1rho: */[1-2-3]/fid
- For Fragments: */??/fid

When we open either set, we will end up with several tens of spectra on our document, what we need to do next is to select all experiments on the page window (just click on the window and do ctrl-A) then press the stack button  , the end result should look something like:



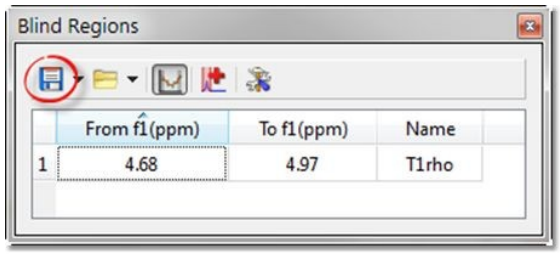
Another alternative would be to use the 'Directory Spectra Stack' script (under the 'Scripts/Import' menu) which will automatically open all the datasets and stack them in one step:



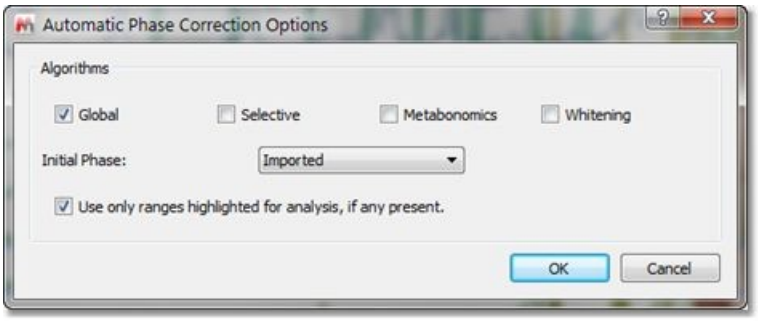
Now we need to increase the intensity of all spectra, to do so we can either use the mouse wheel or the “+” and” –“ keys, the reason performing this step is to be able to monitor the changes that occur when using different processing functions are applied.

For the set of T1rho experiments we are going to use the following reprocessing parameters:

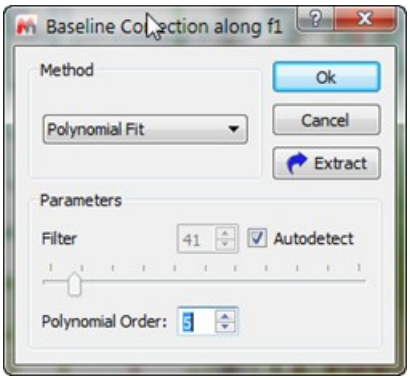
- Blind region between 4.5 and 5.1 ppm, save this blind region, we will use this later:



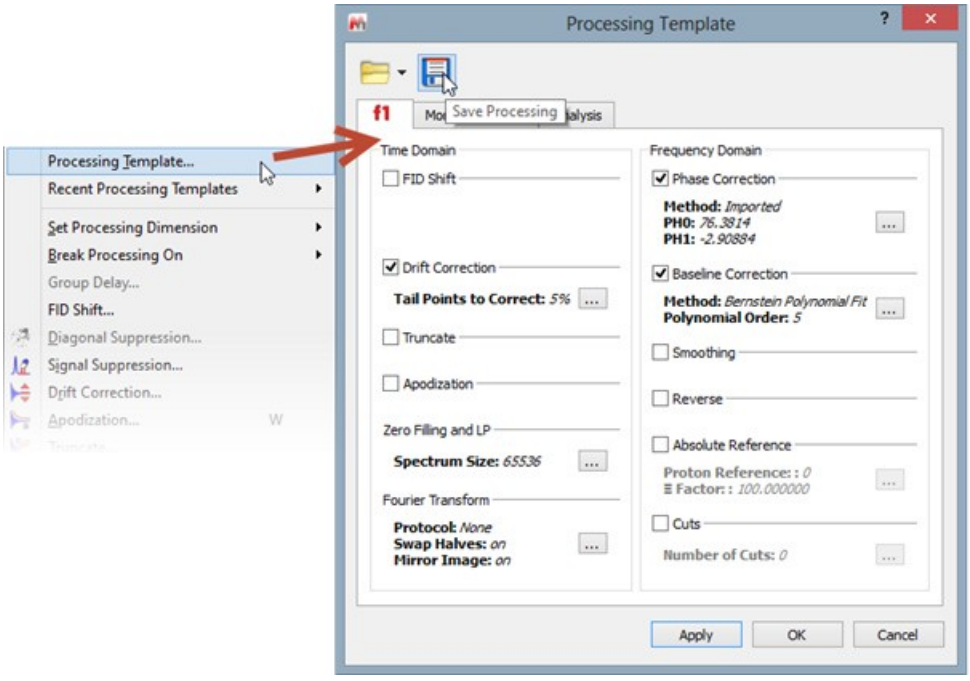
- Use as phase correction algorithm “Global”:



- Set baseline correction to “Polynomial Fit” with order 5:

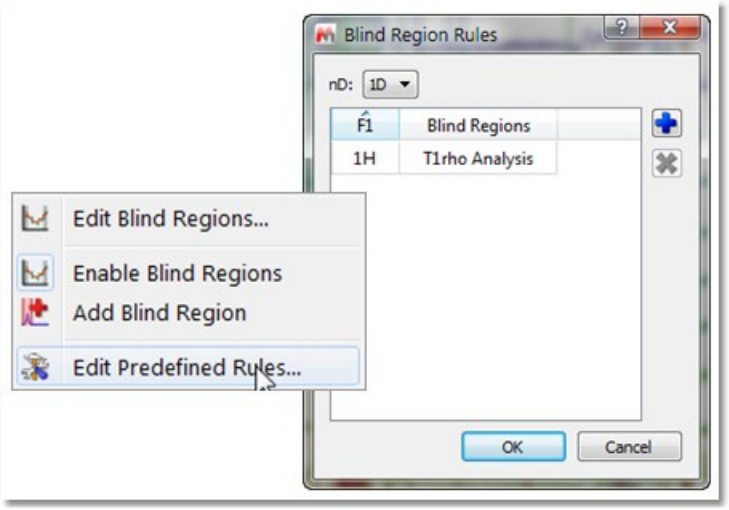


- Save the reprocessing template by going into the menu “Processing/Processing Template”, name the file T1rho Reprocessing



- Last go into 'View/Blind regions/Edit Predefined Rules' and set the previously saved blind region as

you choice for all 1H spectra:

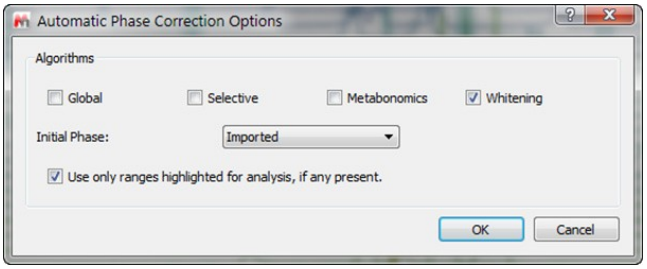


Next, we are going to perform the same procedure as above but now for the spectra of the individual fragments:

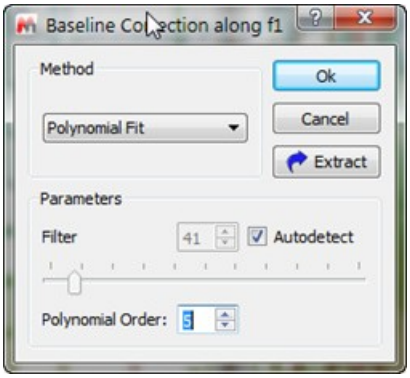
- Import all of the individual spectra using the Multi-Open script with the file mask "*/??/fid"

- Stack the spectra just opened

- Phase correct all spectra using "Whitening" algorithm



- As above perform baseline correction by using the "Polynomial Fit" algorithm, order 5:



- Save this reprocessing template with the name "Fragments_Rep", remember go to

"Processing/Processing Template" menu.

At this stage we should have optimized the reprocessing of both types of experiments.

A few notes on Phase and Baseline Correction

Experiment with the different phase and baseline correction algorithms so the benefits and limitations for each one of them becomes apparent, i.e. Whittaker baseline correction is very aggressive and effective, but note the loss of signal toward the end of the peaks.

Definitions of what each one of these algorithms do and what the "Initial Phase" implies can be found in the relevant parts of the main Mnova NMR manual.
For the case of dealing with the initial settings for phase correction the following recommendations are made:

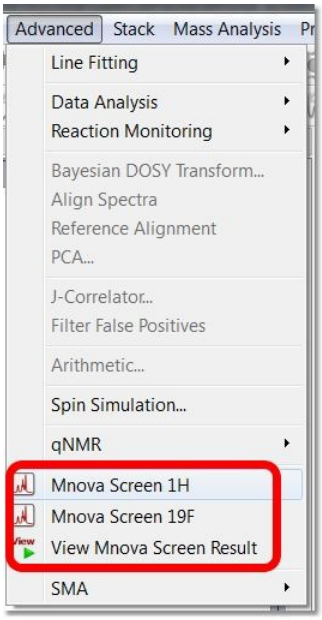
- If no significant phase correction alterations exist proceed to use the Global algorithm with the initial phase set to Imported.
- For data without severe phase distortion due to the fact of the presence of a water suppressed peak we recommend to use a combination of the Global plus Metabonomics
- If the previous recommendation does not produce the desired results, try to define a blind region around the portion of the spectrum which is impossible to phase correctly and tick the option "Use only ranges highlighted for analysis, if any present".
- For spectra which does show phase distortions which are not corrected by any of the other approaches, using only the "Whitening" algorithm is recommended. This works particularly well with wLOGSY type spectra.

Part

4 Data mapping

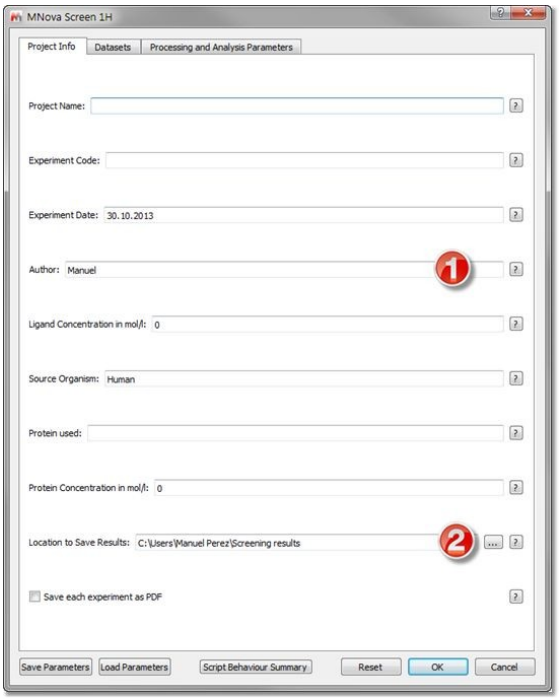
We are now in condition to start the data mapping, the process is analogous to the data importing step we performed above, but with one exception, now we will have to tell the software what every experiment means.

To begin using the screening module, open the 1H NMR Screen from the Advanced Menu:



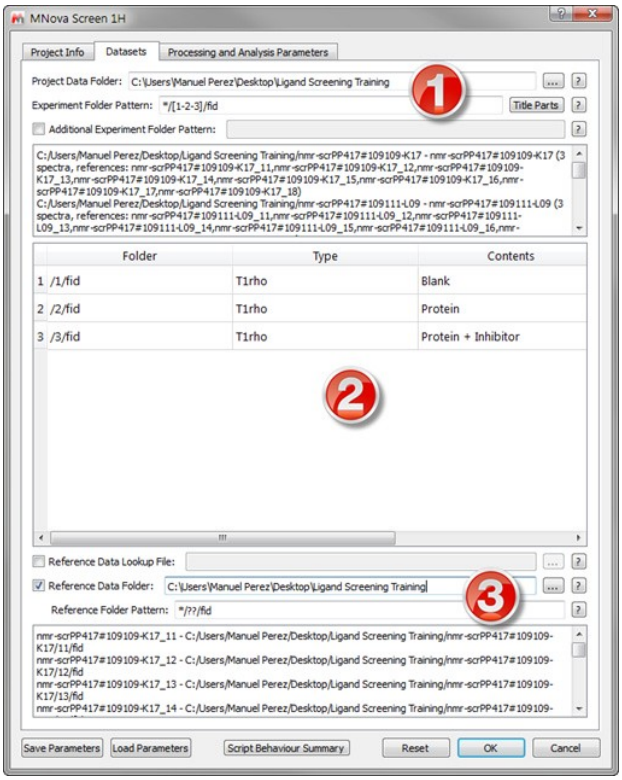
You will be presented with the main Project Info window. We have added detailed descriptions of each one of the fields, what they do how they do it, so at any point when using Mscreen, you can click on any of the question mark icons ? to display these.

The Project Info tab will become the main section where to collect information from the samples, all fields are self-explanatory, but we will stop in order to detail some of those that might require some additional explanation. It is worth mentioning at this stage that the ligand and protein concentration fields are just being used, at the present moment, for informative purposes.



Of importance on this tab are the Author and the Results fields, the former will be used on your results files and the folder containing those experiments will be name with the date/time of the analysis followed by “Author”. The results location is where all the data which is produced by MScreen will be saved. The rest of fields are at the present moment only used for information purposes.

Next we will need to map our data and indicating the software exactly what each experiment is, in order to start doing so move to the “Datasets” tab:

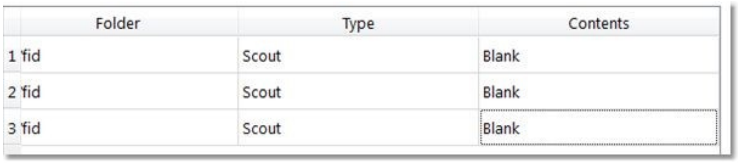


There are three distinctive regions, the Project data region, where we will be mapping the screening experiments, the experiment region where we will indicate the software the nature of each experiment as well as the contents of the sample and finally the Reference data region where we will map each one of the individual fragments.

In this current example we will use a direct way of mapping the data but there are additional methods to do so, which will be explained further on.

The first of steps will be to map the Project Data, in our case this will consist of three different T1rh experiments, in order to do so we will follow these steps:

- Navigate to the directory where the data is saved by clicking on 
- On the “Experiment Folder Pattern” we will use the following path mask “*/[1-2-3]/fid “, without the commas, you should now see this on your experiment window:

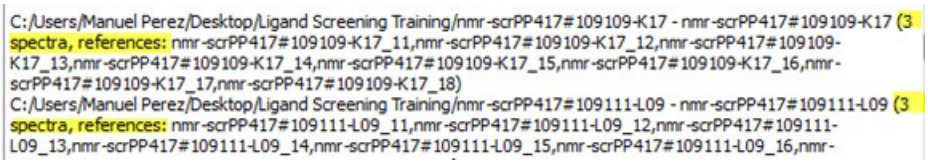


	Folder	Type	Contents
1	fid	Scout	Blank
2	fid	Scout	Blank
3	fid	Scout	Blank

We will need to change the Types and Contents to the correct ones, so they will look something like:

	Folder	Type	Contents
1	fid	T1rho	Blank
2	fid	T1rho	Protein
3	fid	T1rho	Protein + Inhibitor

We shall now proceed at making our references, in order to do so, do navigate to the reference data folder by clicking  , in this case we will use the following file mask “*/??/fid”. After mapping the reference data we should see a scene identical to the one depicted on the 'Datasets' tab, where all the references have been associated to the corresponding experiments, as we can see in better detail on the screenshot below:



Part V

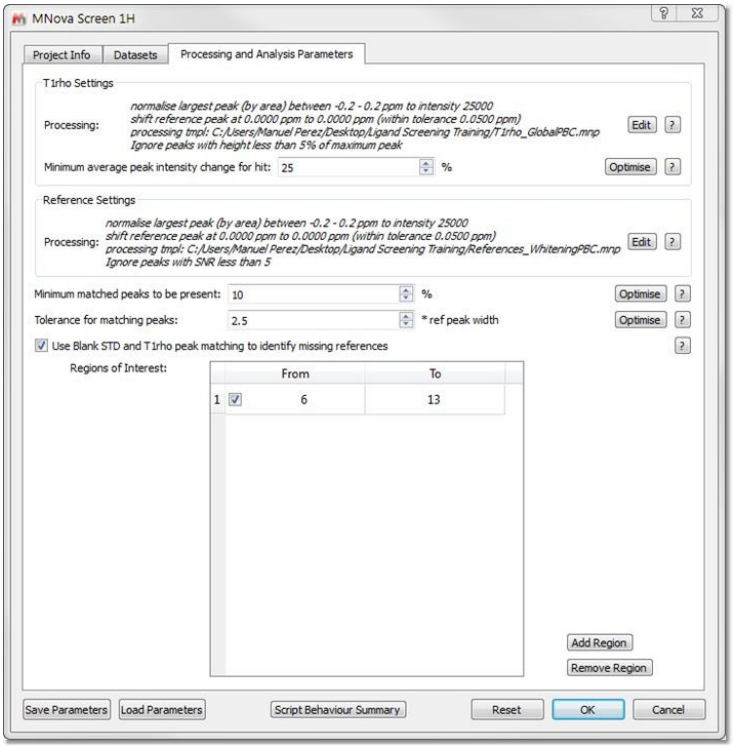
5 Data Processing and Analysis

The next objective will be to set optimum parameters for the reprocessing and analysis of our screening data, we will be only detailing the conditions for T1rho type experiments.

In order to set the scene the changes that will be monitoring are those occurring on the values of the integrals for different peaks. We will not be detailing in this document the theory behind the interactions taking place, for a thorough description please refer to "*Progress in Nuclear Magnetic Resonance Spectroscopy* **44** (2004) 225–256".

At this stage is worth mentioning that for every option dialogs detailing what each one achieves has been created and you can access this by clicking on the question mark ? buttons.

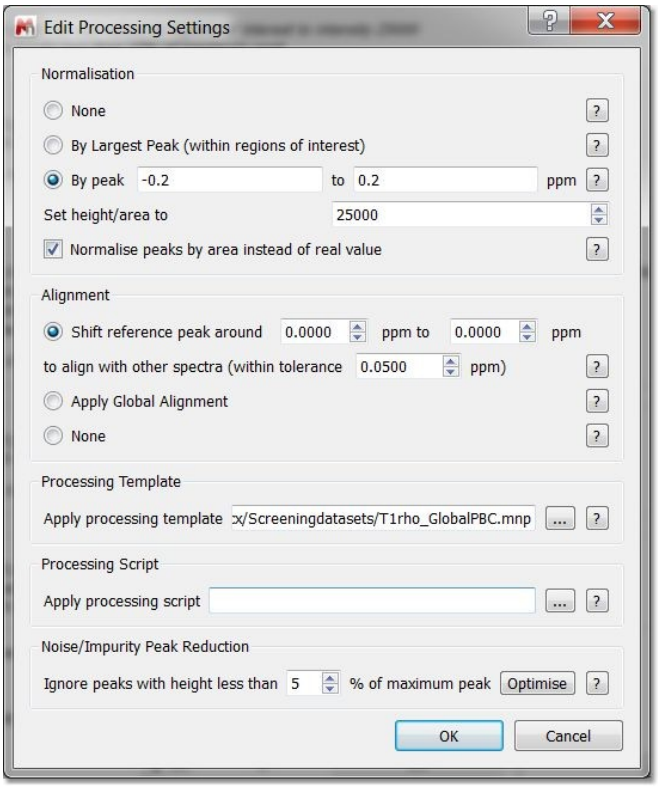
Initially we will be presented with the tab, which will look like:



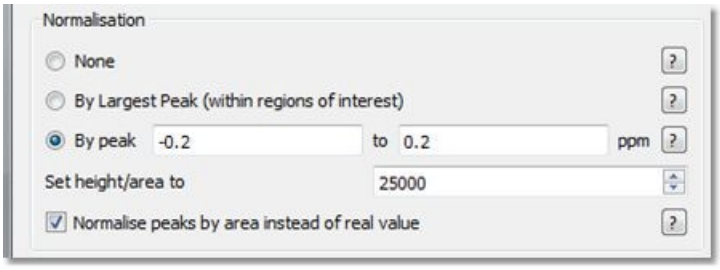
We will start by defining when we would like a peak to be considered a hit, this can be set under:

This value shown in the screenshot above will much depend on the system under study as well as the experience of the scientist, for the case of T1rho experiments setting a 25% as the condition for a hit is a good compromise (please note the "Optimise" button which can be used in this case to produce any number of analysis with different area percentage change conditions).

We need to set the reprocessing parameters for the T1rho experiments, in order to start doing so click on the "Edit" button and the following panel will be presented in front of us:

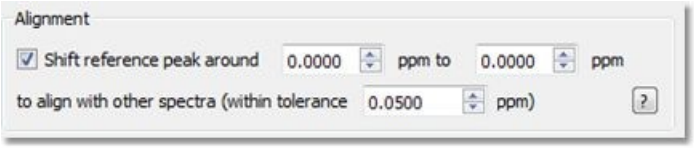


First we are going to set the normalization factor, in our case we will use the TSP peak around 0.0 ppm region – what this is going to enable is a consistent way to compare the changes that occur between the different experiments. The options we will use will be to normalise peaks by area and to use those peaks between -0.2 and 0.2 ppm.



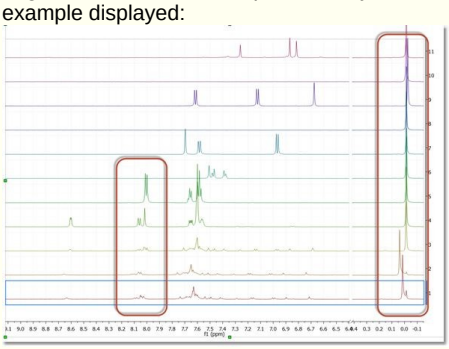
Normalisation of the spectra is an important step and although in some cases could be avoided (if the experimental conditions, including number of scans and receiver gain are identical) the recommendation is to either have an experimental signal, such as TSP, DSS, residual DMSO peak, or an artificially generated peak, that could be used as reference for the normalization. The other options for normalization enable to use the biggest peak within a region of interest as the reference for the normalization step.
The option to normalize peaks by area instead of real value allows to use the intensity of a peak instead of the area to check the behavior of that particular fragment, this is required when analysis wLOGSY experiments in particular

Next we will proceed to define the alignment parameters as shown below:

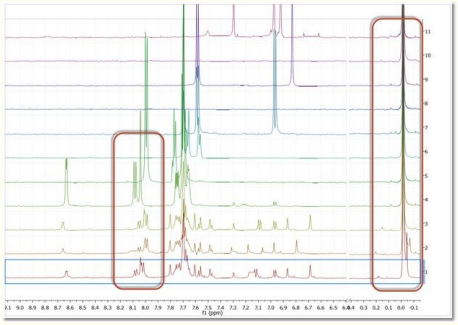


It is worth mentioning at this stage that most of the failures on the automatic analysis arise from the lack of good alignment in between samples, so consider this when examining data.

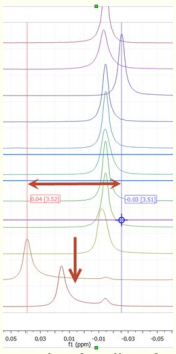
Alignment is a crucial step, the analysis will break down if the peaks are not correctly aligned, see the real example displayed:



It is possible using the alignment options within MScreen to solve these situations, the results can be found on the screenshot below:



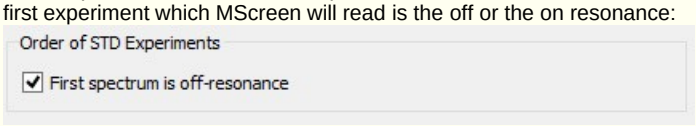
The alignment functionality can be also be found under the Mnova “Advanced” menu. The alignment implementation within Mnova combines a detection of zones using derivatives.
Setting the right limits for the initial alignment is of importance, since having a range which is too wide could allow for incorrect identification of a peak and using this for the alignment step, for example in the following data:



For the family of experiments above we would to set the following alignment parameters: to look for peaks close to 0.01 ppm, shift this to 0.0 ppm and allow a tolerance of 0.02 ppm. With this set of parameters we will cover all of the area of interest.

In the “Processing Template” dialog navigate to the reprocessing template which was created early on.

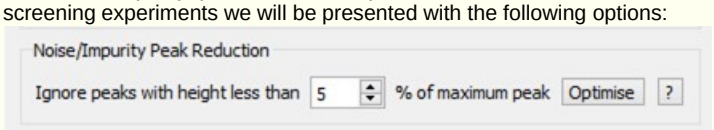
The processing template can be different for each experiment type.
For the particular case of STD experiments we should indicate in the corresponding reprocessing window if the first experiment which MScreen will read is the off or the on resonance:



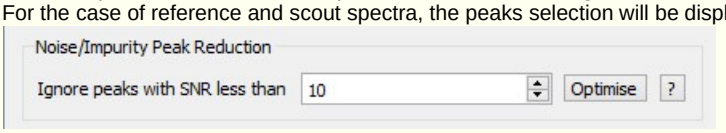
The Processing script option will let the use of specific scripts to be used to call for different reprocessing and reporting functions.

The Noise/Impurity peak reduction will greatly depend on the quality of your data, but using a 5% of maximum peak is a reasonable general value.

The Noise/Impurity peak reduction option comes in two different forms, depending on the sample type, for screening experiments we will be presented with the following options:

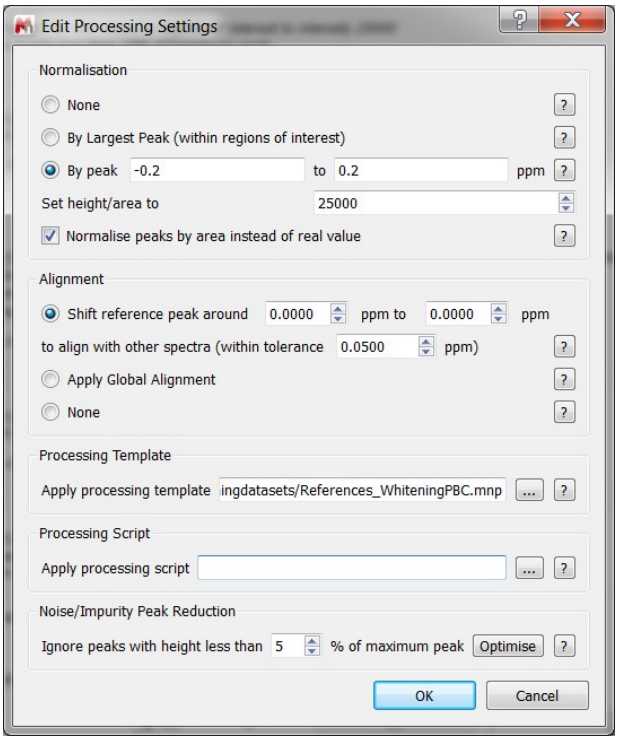


This indicates that those peaks below the marked percentage will not be selected for the analysis, not that the reference peak is the most intense peak within the marked regions of interest and not the total of the spectrum.
For the case of reference and scout spectra, the peaks selection will be displayed as:



The SNR is calculated for the entire of the spectrum and the indicated value will be used as the threshold to pick those peaks which will be used for the analysis.

The settings for reference spectra will be identical to those we set for T1rho experiments, with the exception of the specific reprocessing template, hence make sure that your options look something like:



We will set next the “Minimum matched peaks to be present”, what this option does is to define how many peaks the software needs to match in the mixture to mark a reference sample as present in solution. This option is very useful when we an individual fragment has not been solvated and it is not in the mixture. We will set up this value at 10%:

Next the Tolerance for matching peaks will be set, this particular value, this will be multiplied by the width at half eight of each peak and this will be the interval to the right and left of every peak which will be used to find a matching peak. Note that the higher we make this value the higher we will make the chance of mismatching a set of peaks, and the lower it is the more peaks we might miss in the matching process.

The option of “Use Blank STD and T1rho peak matching to identify missing references” will have to be ticked in this particular example since for this dataset a scout experiment was not acquired.

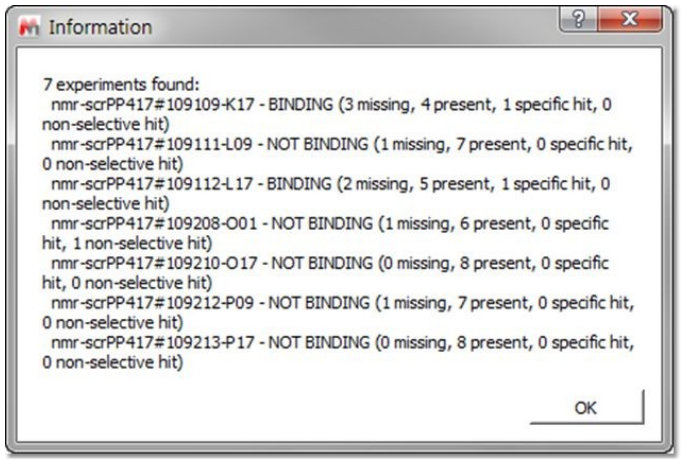
We have to define the spectral areas over which we wish to focus the automatic analysis, you can add as many areas as you wish, always being careful that there are not peaks that could interfere with the analysis (e.g. buffer or residual solvent peaks). In this particular dataset we will focus on the aromatic region between 6 and 13 ppm.

Finally do not forget to save your parameters, check the “Script behavior summary” and click ok to get the results.

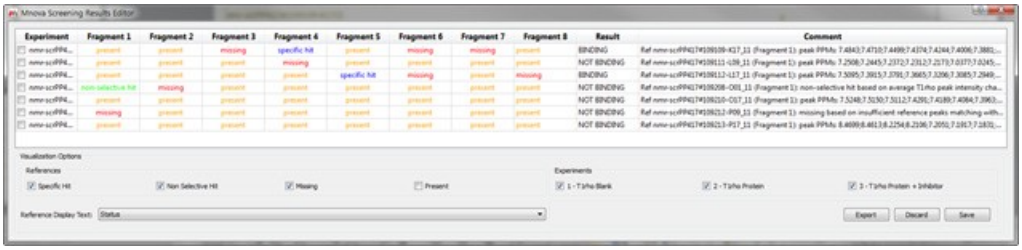
Part

6 Results

Upon completion of the analysis we will be presented with a brief synopsis of the findings from the software:



Press OK and the results viewer should now be in front of you, it should look something like:



If we wish to review the results for any particular experiment we just need to double click on the experiment and the results will be brought on the main window. Before continuing it is worth detailing the four different states that can be attributed to one sample. The color of the individual fragments have meaning, which is as follows:

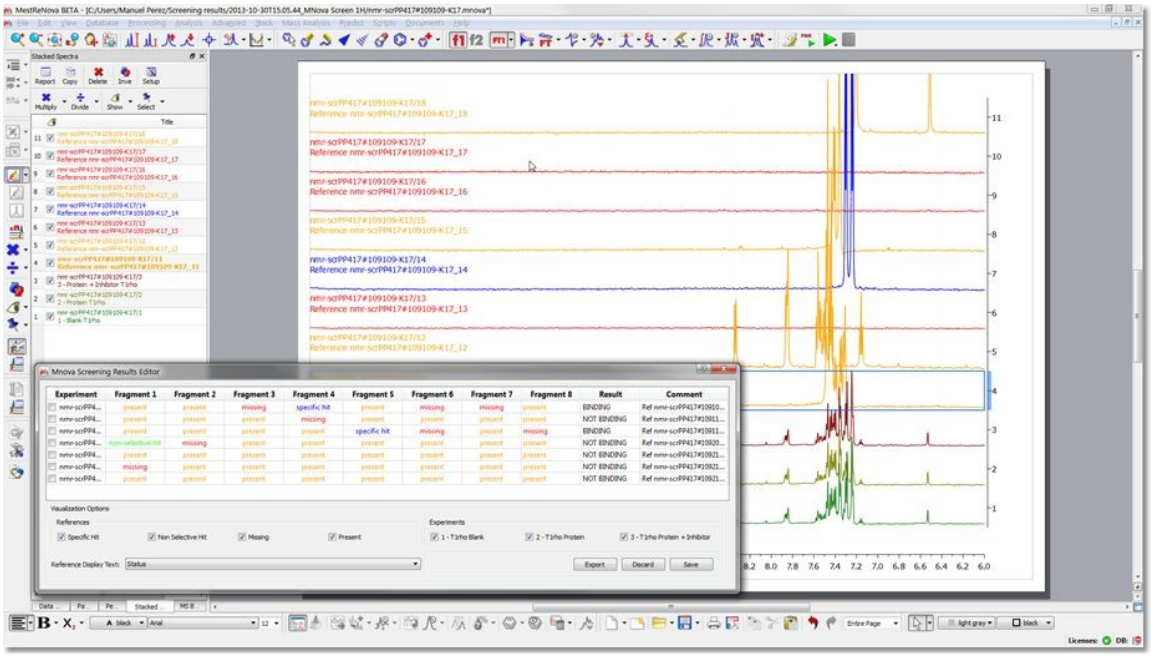
- **Present**: Indicates that the compound has been found but it does not interact with the protein
- **Missing**: no compound signals were found in the region of interest
- **Specific hit**: the compound shows interaction in the sample which contains the protein but not on the sample with protein and competitor
- **Non-selective hit**: the compound shows interaction in both the sample which contains the protein and on the sample with protein and competitor

By clicking on the main interface on the cell corresponding to one specific fragment, it is possible to cycle through each one of the fragment states, change this and later save the results.

The results column summarises which samples have been found to contain binders

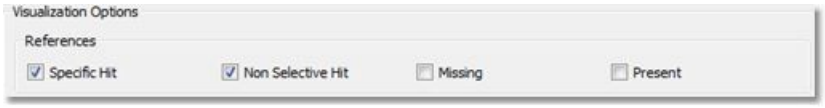
The comment column offers additional information about what peaks have been used for the analysis and which peaks are the ones responsible for marking a hit. In the event of a manual reinterpretation of the spectra, comments can be added to this window, these will be saved.

If we click on the first of experiments we should see something like:



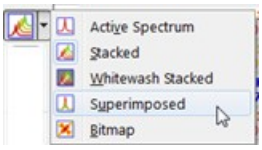
Trick: if you would like to manipulate individual spectra you can do so through the "Stacked Spectra Table", this is accessible from the "View/Tables" menu.

Now we can start reviewing the data, the best way to do so is to focus on those samples which show hits or non-specific hits. To do so, just click on one of the samples, first one for examples and from the screening results editor untick present and missing:

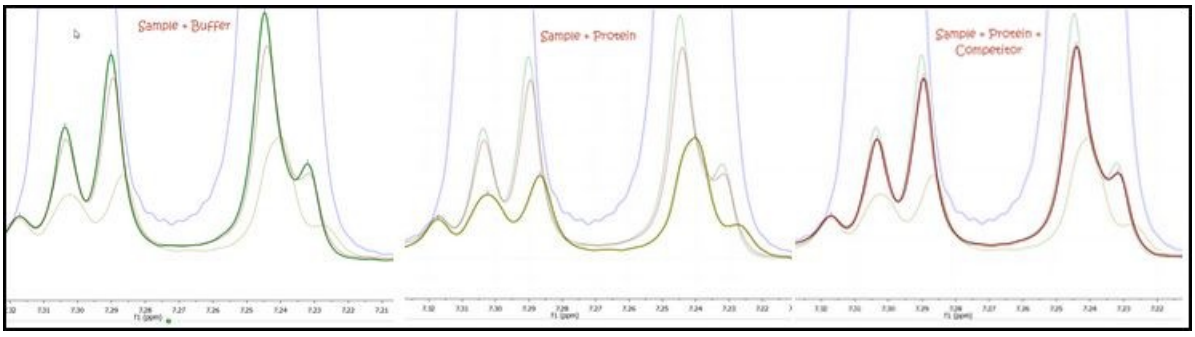


This together with the table of Stacked spectra (available through the view menu or by right-clicking on the toolbar) enable very efficient visual inspection of results.

Superimpose the spectra by going into the stacked button:



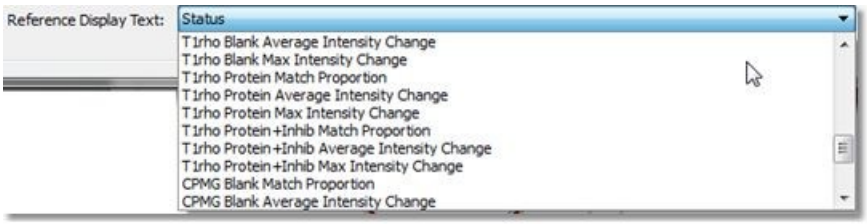
In this manner we can very quickly review the hits, for example for the first of samples we can see by double clicking through the stacked spectra table:



The effect that we can see is how the intensity of the signals of interest for this particular fragment decrease in the presence of the protein under study and how when adding a competitor these recover, fulfilling the conditions for a selective hit.

The 'Experiment visualisation' section: as in the previous by ticking or un-ticking these boxes the view of any of those specific experiments is enabled or disabled/:

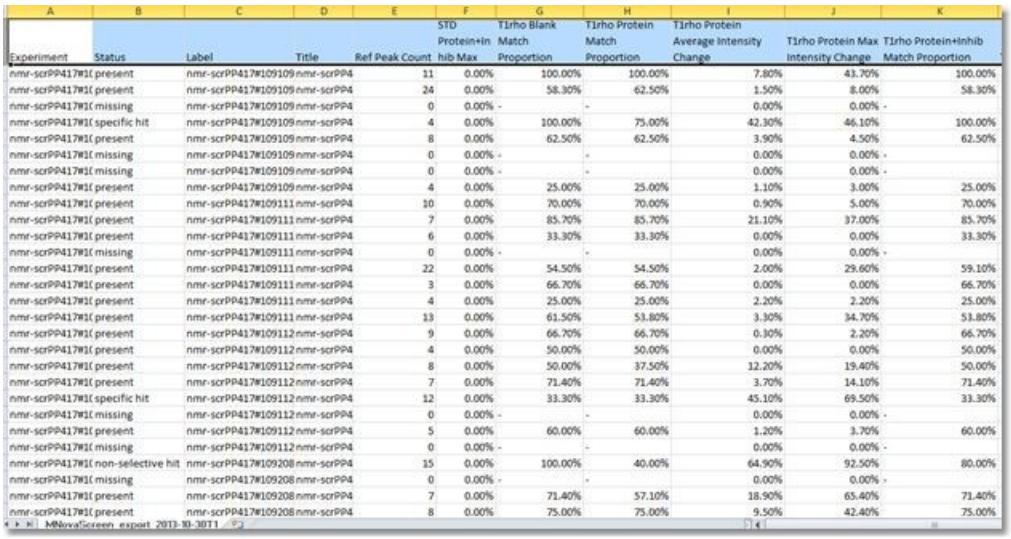
We can also use the results editor to examine the results in a more quantitative manner by using the pull-down menu “Reference Display Text” and selecting the appropriate metric:



For example display T1rho Protein Average Intensity Change:

Experiment	Fragment 1	Fragment 2	Fragment 3	Fragment 4	Fragment 5	Fragment 6	Fragment 7	Fragment 8
nmr-scrPP4...	7.8%	1.5%	0.0%	42.3%	3.9%	0.0%	0.0%	1.1%
nmr-scrPP4...	0.9%	21.1%	0.0%	0.0%	2.0%	0.0%	2.2%	3.3%
nmr-scrPP4...	0.3%	0.0%	12.2%	3.7%	45.1%	0.0%	1.2%	0.0%
nmr-scrPP4...	64.9%	0.0%	18.9%	9.5%	10.4%	3.5%	13.5%	3.1%
nmr-scrPP4...	6.7%	1.8%	4.6%	1.2%	14.0%	4.2%	6.2%	7.6%
nmr-scrPP4...	0.0%	1.4%	3.9%	3.7%	17.3%	6.4%	1.5%	12.7%
nmr-scrPP4...	4.5%	1.4%	2.1%	0.0%	0.5%	5.4%	0.1%	14.4%

In this manner we can notice those fragments that might be close at being considered hits under our criteria. Furthermore, it is possible to create full reports with a combination of all of the metrics that MScreen calculates and export these into .csv format by pressing the “Export” button, after doing so and opening for example on Excel we should see something like:



Experiment	Status	Label	Title	Ref Peak Count	STD Protein+Inhib Max	T1rho Blank Match Proportion	T1rho Protein Match Proportion	T1rho Protein Average Intensity Change	T1rho Protein Max Intensity Change	T1rho Protein+Inhib Match Proportion
nmr-scrPP417#1(	present	nmr-scrPP417#109109	nmr-scrPP4	11	0.00%	100.00%	100.00%	7.80%	43.70%	100.00%
nmr-scrPP417#1(	present	nmr-scrPP417#109109	nmr-scrPP4	24	0.00%	58.30%	62.50%	1.50%	8.00%	58.30%
nmr-scrPP417#1(	missing	nmr-scrPP417#109109	nmr-scrPP4	0	0.00%	-	-	0.00%	0.00%	-
nmr-scrPP417#1(	specific hit	nmr-scrPP417#109109	nmr-scrPP4	4	0.00%	100.00%	75.00%	42.30%	46.10%	100.00%
nmr-scrPP417#1(	present	nmr-scrPP417#109109	nmr-scrPP4	8	0.00%	62.50%	62.50%	3.90%	4.50%	62.50%
nmr-scrPP417#1(	missing	nmr-scrPP417#109109	nmr-scrPP4	0	0.00%	-	-	0.00%	0.00%	-
nmr-scrPP417#1(	missing	nmr-scrPP417#109109	nmr-scrPP4	0	0.00%	-	-	0.00%	0.00%	-
nmr-scrPP417#1(	present	nmr-scrPP417#109109	nmr-scrPP4	4	0.00%	25.00%	25.00%	1.10%	3.00%	25.00%
nmr-scrPP417#1(	present	nmr-scrPP417#109111	nmr-scrPP4	10	0.00%	70.00%	70.00%	0.90%	5.00%	70.00%
nmr-scrPP417#1(	present	nmr-scrPP417#109111	nmr-scrPP4	7	0.00%	85.70%	85.70%	21.10%	37.00%	85.70%
nmr-scrPP417#1(	present	nmr-scrPP417#109111	nmr-scrPP4	6	0.00%	33.30%	33.30%	0.00%	0.00%	33.30%
nmr-scrPP417#1(	missing	nmr-scrPP417#109111	nmr-scrPP4	0	0.00%	-	-	0.00%	0.00%	-
nmr-scrPP417#1(	present	nmr-scrPP417#109111	nmr-scrPP4	22	0.00%	54.50%	54.50%	2.00%	29.60%	59.10%
nmr-scrPP417#1(	present	nmr-scrPP417#109111	nmr-scrPP4	3	0.00%	66.70%	66.70%	0.00%	0.00%	66.70%
nmr-scrPP417#1(	present	nmr-scrPP417#109111	nmr-scrPP4	4	0.00%	25.00%	25.00%	2.20%	2.20%	25.00%
nmr-scrPP417#1(	present	nmr-scrPP417#109111	nmr-scrPP4	13	0.00%	61.50%	53.80%	3.30%	34.70%	53.80%
nmr-scrPP417#1(	present	nmr-scrPP417#109112	nmr-scrPP4	9	0.00%	66.70%	66.70%	0.30%	2.20%	66.70%
nmr-scrPP417#1(	present	nmr-scrPP417#109112	nmr-scrPP4	4	0.00%	50.00%	50.00%	0.00%	0.00%	50.00%
nmr-scrPP417#1(	present	nmr-scrPP417#109112	nmr-scrPP4	8	0.00%	50.00%	37.50%	12.20%	19.40%	50.00%
nmr-scrPP417#1(	present	nmr-scrPP417#109112	nmr-scrPP4	7	0.00%	71.40%	71.40%	3.70%	14.10%	71.40%
nmr-scrPP417#1(	specific hit	nmr-scrPP417#109112	nmr-scrPP4	12	0.00%	33.30%	33.30%	45.10%	69.50%	33.30%
nmr-scrPP417#1(	missing	nmr-scrPP417#109112	nmr-scrPP4	0	0.00%	-	-	0.00%	0.00%	-
nmr-scrPP417#1(	present	nmr-scrPP417#109112	nmr-scrPP4	5	0.00%	60.00%	60.00%	1.20%	3.70%	60.00%
nmr-scrPP417#1(	missing	nmr-scrPP417#109112	nmr-scrPP4	0	0.00%	-	-	0.00%	0.00%	-
nmr-scrPP417#1(	non-selective hit	nmr-scrPP417#109208	nmr-scrPP4	15	0.00%	100.00%	40.00%	64.90%	92.50%	80.00%
nmr-scrPP417#1(	missing	nmr-scrPP417#109208	nmr-scrPP4	0	0.00%	-	-	0.00%	0.00%	-
nmr-scrPP417#1(	present	nmr-scrPP417#109208	nmr-scrPP4	7	0.00%	71.40%	57.10%	18.90%	65.40%	71.40%
nmr-scrPP417#1(	present	nmr-scrPP417#109208	nmr-scrPP4	8	0.00%	75.00%	75.00%	9.50%	42.40%	75.00%

This will allow further analysis of the data in alternative ways.

Part

7 Analysis definitions

The experiments considered are detailed in the following table as well as how we set ourselves to perform the analysis for each one of them:

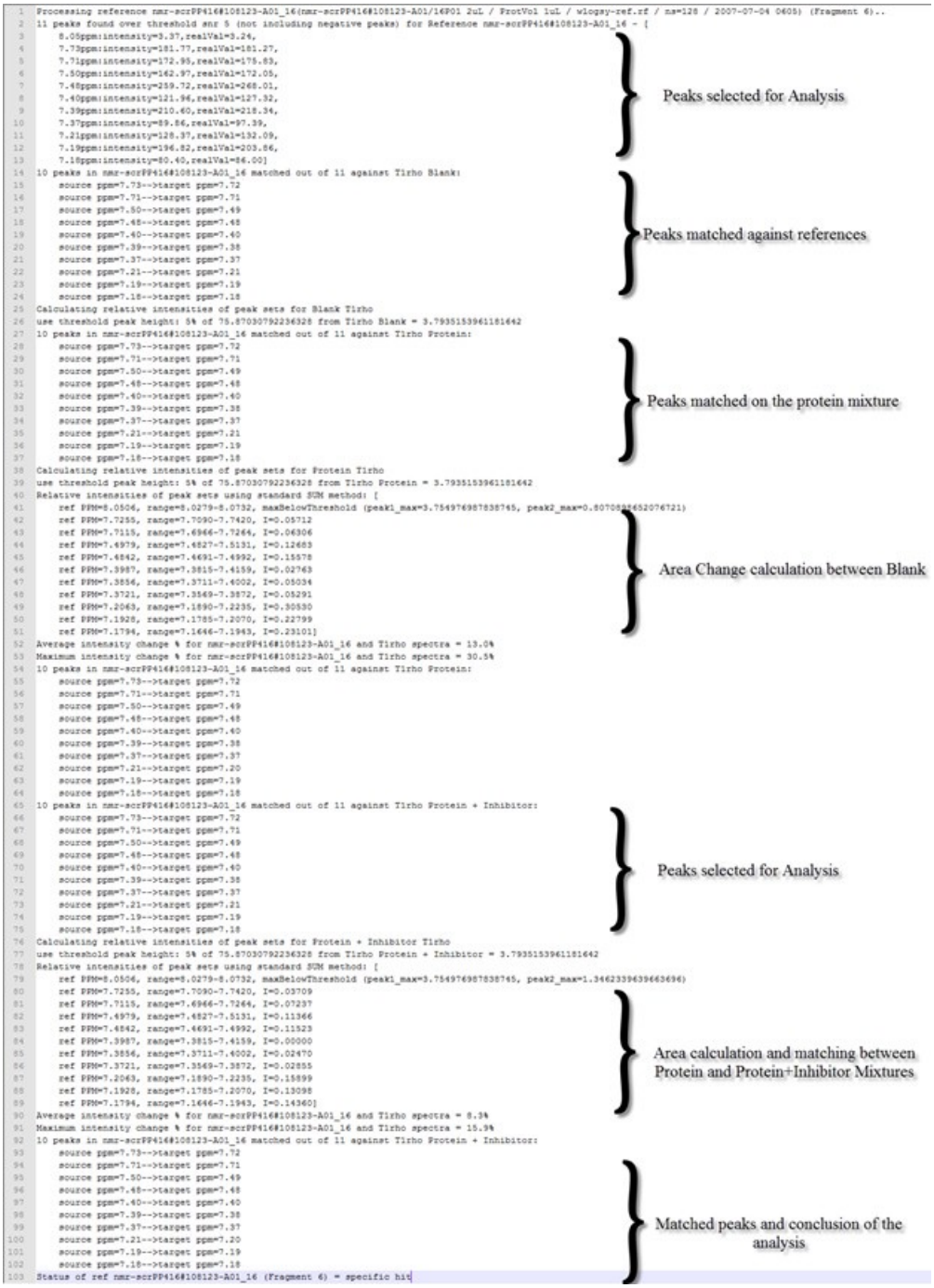
Experiment	Sample contents	What to observe?	How to analyze?
Reference spectra	Single library compound	The reference peaks of the library compounds	Process the spectra, pick ref. peaks by GSD
Saturation transfer Detection (STD)	Mixture of N library compounds and protein. Blank (no protein) and/or competition (with protein and inhibitor) may be used for comparison	The on-resonance spectral peaks are expected to decrease if that compound is binding with the protein.	If the STD difference spectra are used, then the peaks are picked and matched to ref peaks. If the on/off-resonance spectral pairs are used, then the ref peaks are mapped to the STD pair. The change of peak intensities within the mapped regions are calculated and averaged. If available, blank and competition data will be used to detect false positive and non-specific binding if available.
T1ρ (and CPMG)	Mixture of N library compounds and protein. Blank (no protein) and/or competition (with protein and inhibitor) may be used for comparison	The peaks from the long spin-lock time spectrum are expected to decrease compared with the short spin-lock one, if that compound is binding with the protein. If inhibitor is used, the peaks are expected to restore to the original height if it is a specific binding	The ref peaks are mapped to the T1ρ pair. The change of peak intensities within the mapped regions are calculated and averaged. If available, blank and competition data will be used to detect false positive and non-specific binding if available.
WaterLOGSY	Mixture of N library compounds without protein (blank), and with protein. Competition sample (with protein and inhibitor) may be used for comparison	Negative are expected in the blank . Positive peaks are expected for a compound that binds with the protein. If inhibitor is used, such peaks are expected to restore to negative if it is a specific binding	The ref peaks are mapped to the wLogsy spectra. The change of peak intensities within the mapped regions are calculated and averaged. If available, blank and competition data will be used to detect false positive and non-specific binding if available.

Part VIII

8 Log Files Structure

A great deal of information is produced by the automatic analysis of MScreen over one single run and all and each one of the steps are recorded for every sample on the log file.

The structure of this log is quite detailed and all of the decision making process can be followed, for example for one particular fragment this would be the full analysis:



Part

9 19F Analysis

From version 1.1 of MScreen it is also possible to perform analysis over ^{19}F screening data, all of the analysis, data manipulation and reporting capabilities open for ^{1}H are also available for ^{19}F.

There are just a two of points of difference between the ^{1}H and the ^{19}F automatic analysis setting. The first one is the number of peaks which have to be detected to consider an individual fragment to be present. The second one is the tolerance for matching peaks, in the scenario of 19F this is a fixed value in ppm which must be introduced.

Thank you!

Thank you for reading this manual, and for purchasing this release version of Mnova. We will be very keen to read your feedback on the application, to hear about any bugs you may find and to also listen to any additional ideas or suggestions you may have.

Please remember that you can send all those, and any queries about the software, or requests for help, to:

support@mestrelab.com

Keep checking our web site (www.mestrelab.com) for additional information on our range of software packages, and for news on our company.