

How to Use FLYA/CYANA module with Deep Neural Network

1) required files for FLYA calculation

FLYA setup module will help you to generate files required for FLYA calculation. FLYA is a function implemented in CYANA which can be used for fully automated signal assignment and structure calculation. (CYANA ver. 3.75-3.98)

Recommended spectrum data sets for backbone signal assignments;

2D ^1H - ^{15}N HSQC, 3D-HNCO, 3D-HN(CA)CO, 3D-HNCACB, 3D-CBCA(CO)NH

For assignments of all signals including backbone and side-chain atoms (automated structure determination);

3D-HCCH-TOCSY for aliphatic and aromatic

3D- ^{15}N edited NOESY, 3D- ^{13}C edited NOESY for aliphatic and aromatic

It is strongly recommended that you should increase the data points for the NOE or $^{13}\text{C}/^{15}\text{N}$ dimension using zero-filling, linear prediction and non-linear sampling. The followings are the successful examples of spectrum data sets (complex);

3D HCCH-TOCSY X: 2048 Y:256 Z:256

^{15}N -edited NOESY X: 1024 Y: 512 Z:256

^{13}C edited NOESY X: 2048 Y:512 Z:256

For the FLYA calculation, the following files are required;

CACL.cya, int.cya Setting files for FLYA calculation.

*.seq Sequence file

*.peaks Peak table file

For CNN/FNN noise filter

CNTK ver. 2.0 is required. Please download CPU-only version from the Microsoft web-site:

<https://cntk.ai/dl/c-2.0.html>

just download CNTK-2-0-Linux-64bit-CPU-Only.tar.gz

tar xvzf /home/naohiro/opt/CNTK-2-0-Linux-64bit-CPU-Only.tar.gz

Install OpenMPI ver. 1.10.x

<https://www.open-mpi.org/software/ompi/v1.10/>

download and extract the compressed file, then install it. (If you don't know what to do please consult your system administrator)

[Important]

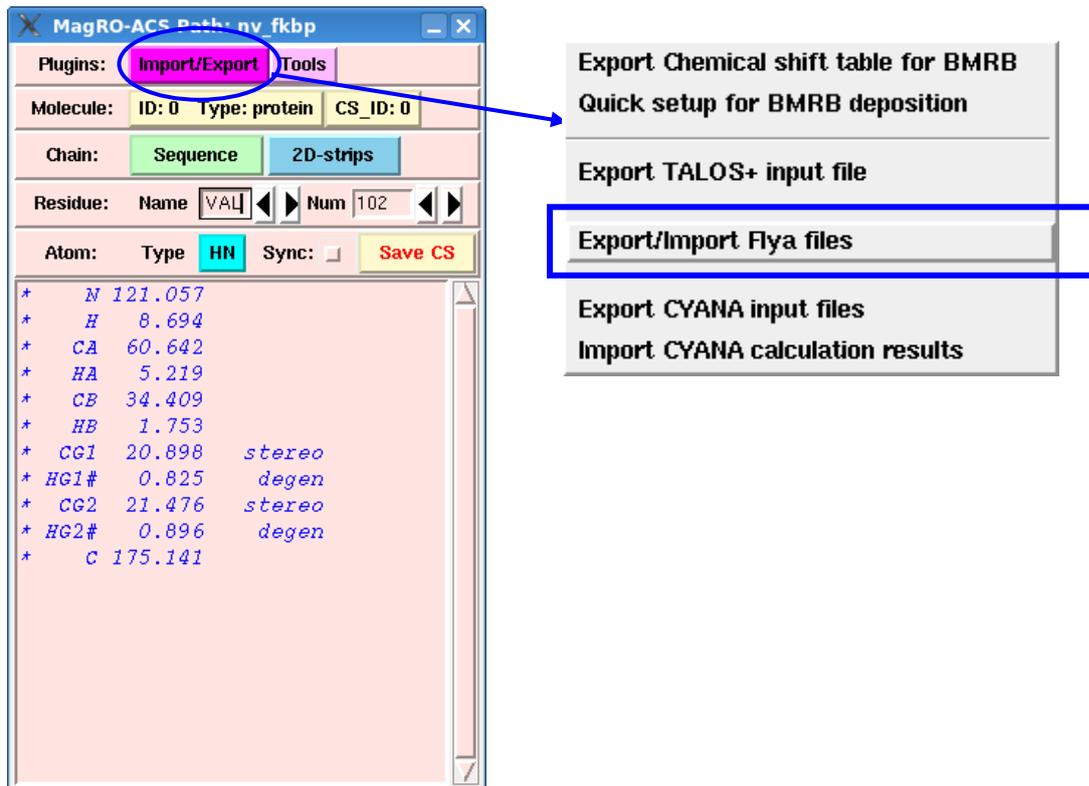
The accuracy of the automated assignments strongly depends on spectrum quality. If you find that the peaks in the spectra are significantly overlapped and signal missing, you will have to reconsider the plan of automated signal assignments.

[Warning]

The peak lists from the spectrum folded (peaks are detected in folded/aliased/frame-shifted positions) may decrease accuracy of automated assignments. However, if you have ^{13}C -HSQC for full region, the program MagRO can recover the folded peaks in HCCH-TOCSY and ^{13}C -edited NOESY to unfolded position.

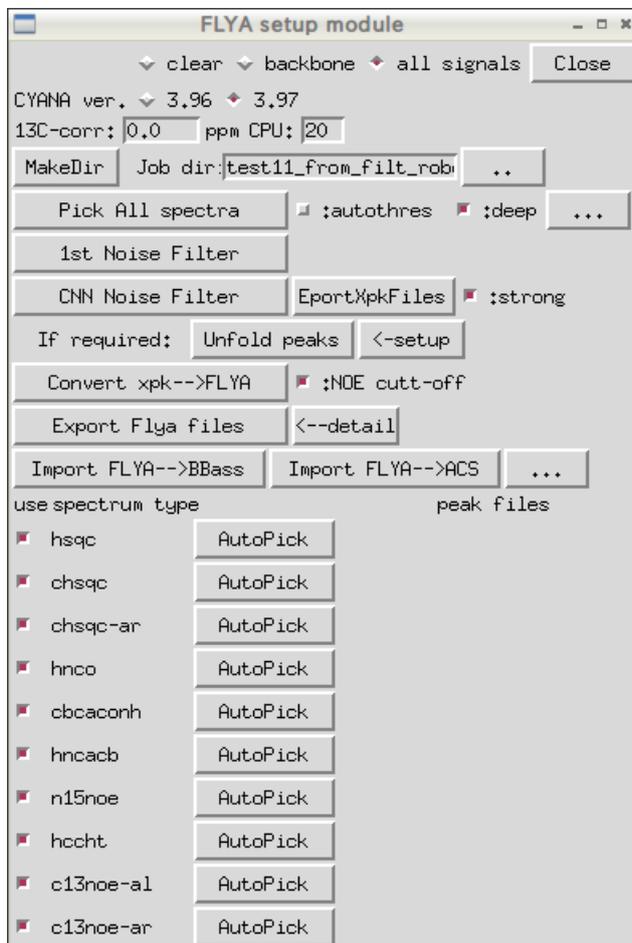
2) Lets start FLYA calculation

Press "Import/Export" button on the Acs module, then select "Export/Import Flya files"



The screenshot shows the 'MagRO-ACS Pathway' window. The 'Import/Export' button is circled in pink. A blue arrow points from this button to a menu of options. The 'Export/Import Flya files' option is highlighted with a blue rectangle. The menu options are:

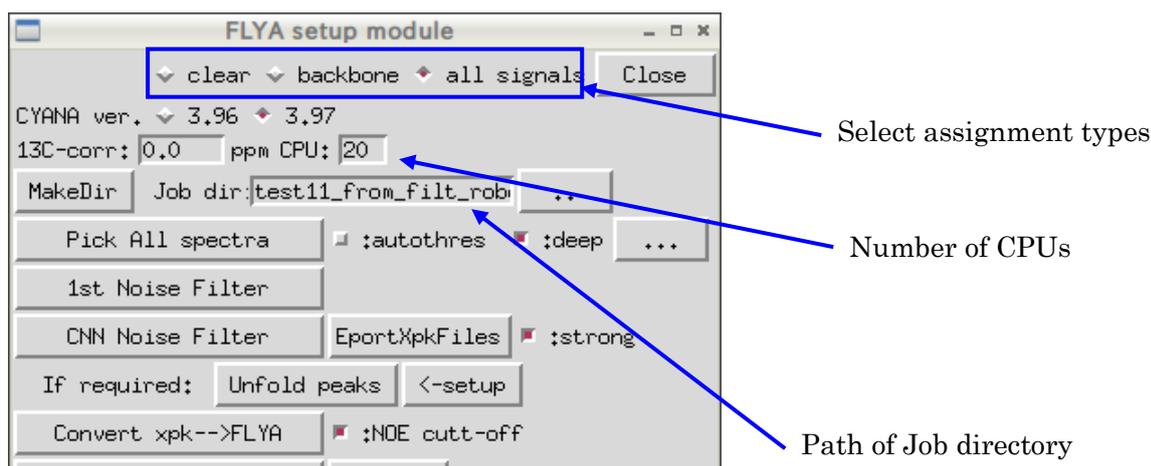
- Export Chemical shift table for BMRB
- Quick setup for BMRB deposition
- Export TALOS+ input file
- Export/Import Flya files**
- Export CYANA input files
- Import CYANA calculation results



The screenshot shows the 'FLYA setup module' window. The window contains various configuration options for NMR data processing. The options are:

- clear
- backbone
- all signals
- Close
- CYANA ver. 3.96
- 13C-corr: 0.0 ppm CPU: 20
- MakeDir Job dir: test11_from_filt_rob ..
- Pick All spectra :autothres :deep ...
- 1st Noise Filter
- CNN Noise Filter EportXpkFiles :strong
- If required: Unfold peaks <-setup
- Convert xpk-->FLYA :NOE cutt-off
- Export Flya files <--detail
- Import FLYA-->BBass Import FLYA-->ACS ...
- use spectrum type peak files
- hsqc AutoPick
- chsqc AutoPick
- chsqc-ar AutoPick
- hnco AutoPick
- cbcaconh AutoPick
- hncacb AutoPick
- n15noe AutoPick
- hccht AutoPick
- c13noe-a1 AutoPick
- c13noe-ar AutoPick

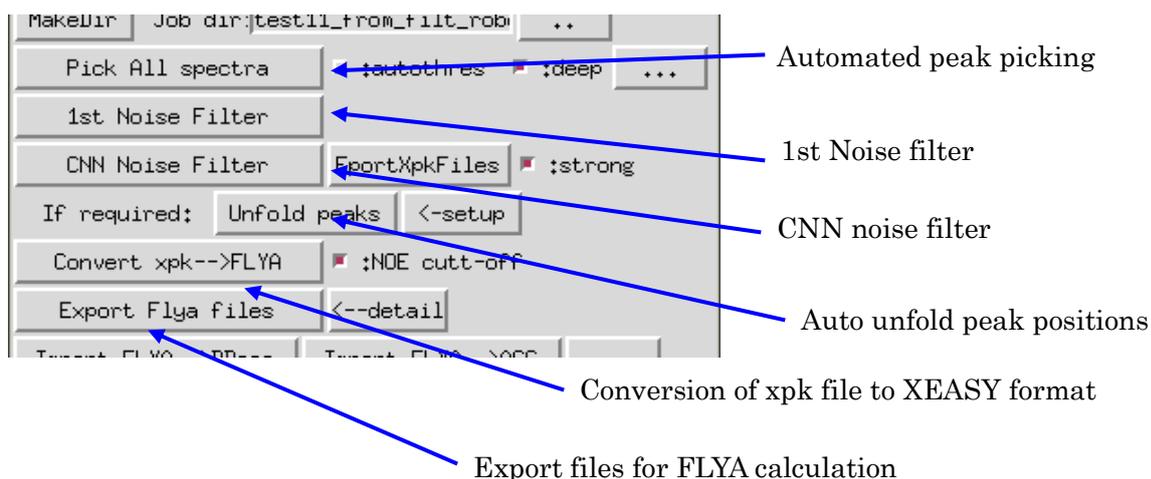
Then you will see the setting up window as shown in the left panel.



a) Firstly, you have to choice type of assignments, backbone signal only or all signals.

And select version of CYANA (FLYA). The default is 3.97 compatible for ver. 3.98.

b) Secondary, set the directory path for FLYA calculation. Directly type the path or get the directory by pressing ".." button. If you would like to create the directory, press "MakeDir" button.



c) Press "Pick All spectra" button to start automated peak picking for all spectra.

The check-button "autothres" means the quick-picker runs automated threshold control mode.

The check-button "deep" means the quick-picker runs automated threshold control mode and will set a little lower threshold.

[Important] If the "autothres: OFF and "deep: OFF", you have to set spectrum threshold properly for each spectrum window.

Press check-button auto-thres to run the quick-picker with automated threshold control mode.

Press check-button deep which means the quick-picker will pick the peaks for deep learning noise filter mode.

d) Press "1st Noise Filter" button to start noise filtration for all peak tables.

Noise filter function will use peak table derived from 2D ¹⁵N-HSQC and ¹³C-HSQC for aliphatic and aromatic as mask table.

e) CNN Noise filter **[New function]**

This is the new function of MagRO using deep neural networks. User has to install CNTK ver. 2.0 and OpenMPI ver. 1.10.x. CNTK has to run as “cntk” command in your Linux system. Please consult your system administrator if you do not know what to do.

This function need at least 4-6GB memory and 2-4-CPU (2.0-30GHz). It will take 20-30min to prepare image data for CNTK recognition. After the automated filtration jobs you will find *_cnn.xpk files in the job directory. The check-button “strong” means the power level of CNN filtration. Press “ExportxpkFiles” you can run the final stage of CNN filtration jobs without image generation (It usually takes a few seconds).

[Important] The capability of the noise elimination power is around 95-99.0% depends on the spectrum quality. Also user has to be careful, this filter eliminates a little number of real peaks (~1%). FLYA and CYANA have a high robustness of a few percent of noise and missing signals.

f) You can refold peaks in folded/aliased/frame-shifted position by pressing "Unfold peaks" button.

[Important] To use this function, you need to prepare spectra for 2D ¹³C-HSQC for aliphatic and aromatic regions (chsqc and chsqc-ar) which have been acquired without spectrum folding in ¹³C dimension. After the automated unfolding jobs you will find *_unfold.xpk files in the job directory.

g) Conversion of xpk files to CYANA (XEASY) peak format

To convert the peak table files for FLYA calculation, press "Convert xpk->FLYA" button.

The check-button “NOE cut-off” means the convertor will filter weak peaks whose intensity is smaller than the 2.0 x threshold value.

The priority is set for the conversion based on the following file name:

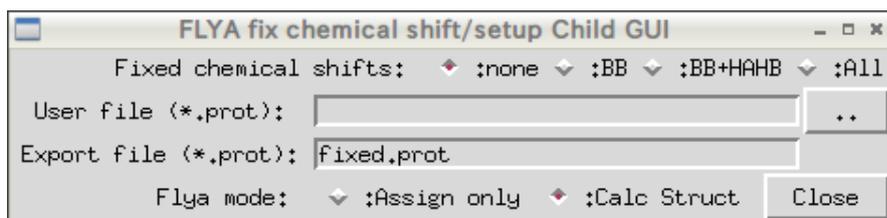
*_auto.xpk -> *_filt.xpk -> *_cnn.xpk -> *_unfold.xpk -> *_refine.xpk

If you correct or edit the peak lists manually, you have to store the peak lists as *_refine.xpk so that the auto conversion will use the files with highest priority.

h) Setup FLAYA calculation

Press “Export FLYA files”, the program automatically prepare file for FLYA calculations such as CALC.cya, init.cya, sequence file (as protein.seq).

If you want to run FLYA for fully-automated signal assignments and structure calculation, press “detail” then check “Calc struct”.



3) Detail of the files created by FLYA setup module

CLAC.cya,	A macro file for FLYA calculation. You can describe spectrum names, structure file name
init.cya	Settings for FLYA calculations such as CPU number, sequence file name, library file name, etc.
protein.seq	Sequence of target protein. Should be described by three letter codes
***_auto.xpk	Peak table identified by auto peak picking
***_filt.xpk	Peak table filtered by 1st Noise filter command
***_cnn.xpk	CNN/FNN filtered peaks
***_unfold.xpk	Peak table with unfolded peaks
***.peaks	peak table in XEASY format

[Important] If you edit xpk file, you should overwrite *_refine.xpk as mentioned above. For example, hncacb-al_refine.xpk.**

4) Execute FLYA calculation

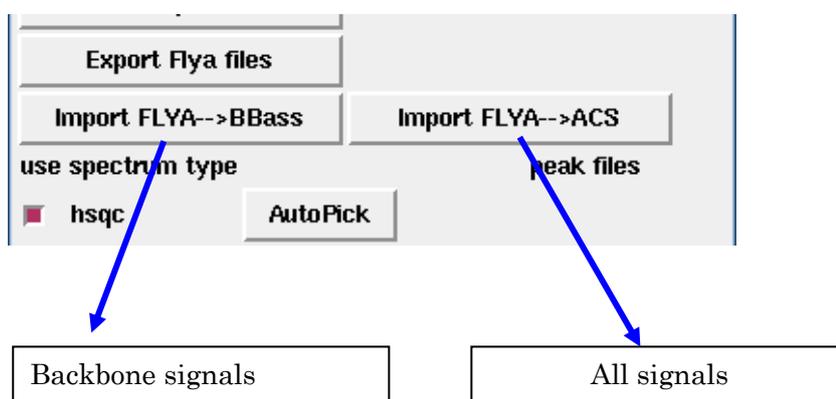
You can start FLYA calculation on the job directory using following command;

```
/opt/cyana3.97/cyana CALC.cya
```

The calculation time using 4CPUs is 5~10min for backbone signals and 60-90min for all signals.

5) Import results of FLYA calculation

After the FLYA calculation you will find an assigned chemical shift table as flya.tab.



In this table the most important parameter is "inside" on the 5th column. The value estimate the frequency of the assigned chemical shift value in the calculation. The FLYA module will pick up the assigned chemical shifts with the value above 80% (labeled with strong) as they are trustful enough.

Total number of shift values: 32057

Cutoff for extent : 16.00

Atom	Residue	Ref	Shift	Dev	Extent	inside	inref	
N	ARG	3	116.353		20.0	55.7	0.0	
H	ARG	3	8.181		20.0	54.3	0.0	
CA	ARG	3	59.572		20.0	99.8	0.0	strong
HA	ARG	3	3.795		20.0	99.7	0.0	strong
CB	ARG	3	29.085		20.0	99.7	0.0	strong

To import the results of FLYA calculation for backbone signals, press "Import FLYA->BBass" button and for all signals press "Import FLYA-->ACS" button.

After the import, the program will overwrite the assigned chemical shifts in assign_NN.txt file for backbone signals and ACS table MagDB/protein_0_0_acs.db file for all signals, respectively.

[Important] The assigned chemical shifts in assign_NN.txt and MagDB/protein_0_0_acs.db will be overwritten with imported values. Please make sure to take backup of the files before do this job.