Kujira ver. 0.99x instruction manual

0. Contents 0-1

Setting up Kujira

1. Installation

You have three things to do for the installation:

> 1. Installation of NMRView ver5.x C-program version
> 2. Obtain the sequence file of protein as *.seq file
>   (Amino acid sequence of sample written in 1- or 3-letter code.)
> 3. Get the kujira_v0.9XX.tar.gz file

The increment in the first digit of the version number means that a minor revision has been made in the package. Assuming that now you have installed NMRView in the directory "/home/yourname/nv_peptide";

1) Make "matrix" directory;

        mkdir /home/yourname/nv_peptide/matrix

The matrix directory will be your working directory of Kujira.   Place all the spectrum *.nv, sequence *.seq files and peak table *.xpk files in the matrix directory.

2) Uncompress and copy kujira to somewhere in your home directories.

        cp -r kujira_v0.9xx   ~/bin/kujira

3) Then move to ~/bin/kujira, you will find a directory "cshfiles".   Move it to upper directory:

        cd ~/bin/kujira
        mv cshfiles ../
        cd ../cshfiles

There is a template script file, kj1.csh.

3) Edit the kj1.csh, with a text_editor, then modify the following line as the script specifies the nv_peptide:

        [BEFORE]   setenv NMRVIEW5HOME ~/nv_XXX
        [AFTER]   setenv NMRVIEW5HOME ~/nv_peptide

4) Go back to your home directory, edit .cshrc file.   Add the following lines;

         alias kujira '~/bin/kujira/start_kujira.tcl $NMRVIEW5HOME'
         alias kj1 'source ~/bin/cshfiles/kj1.csh'

5) To activate the edited settings, execute 'source' command:

        source ~/.cshrc

If you do not see any error message, the installation has been finished.

To start Kujira, type "kj1".

2. Update of Kujira

To update Kujira, what you have to do is just replace the old version of kujira with the latest one.

If the update has been successfully finished, you will find version name on the right corner of the startup module,

or on the left corner of the skyblue module,

3. Initial setup of Kujira

The first thing to do is making the Acs directory which will manage the sequence and chemical shifts of your NMR sample.
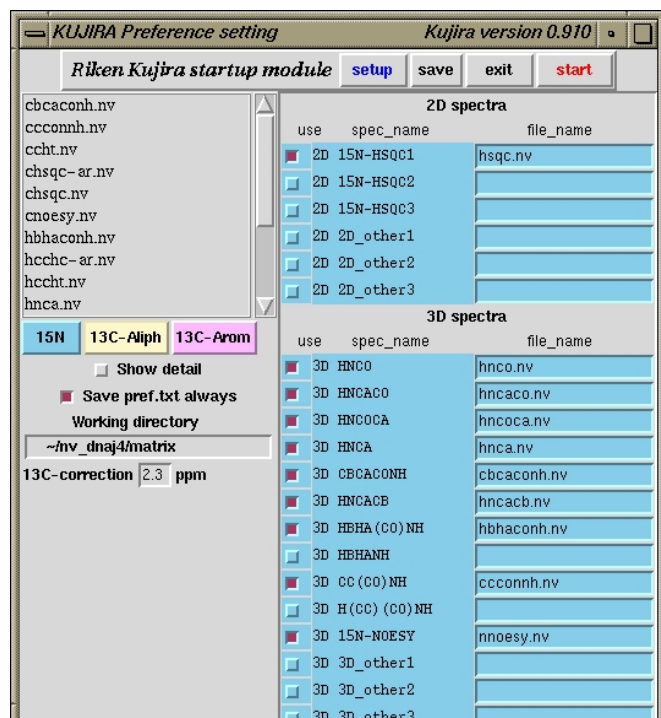
If you start Kujira, you will see the startup module as shown in the left below panel.

If the startup of Kujira is very first time, kujira will give you an error message telling you, "Could not find Acs directory" and Kujira will ask you to make Acs directory. If you answer 'Yes', you will see a text editor as follows:
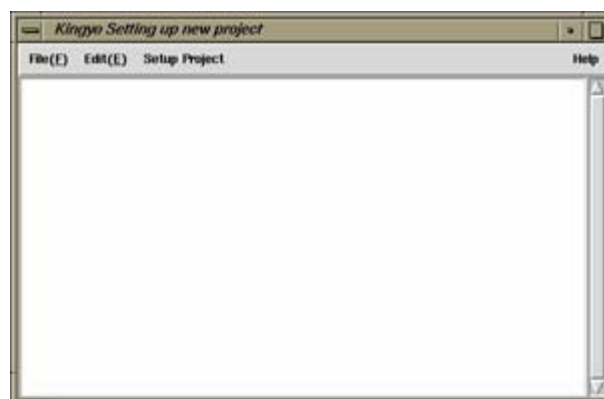
On this editor, you can use 'Cut' , 'Copy' and 'Paste' functions like the other text_editor as well as 'open file', 'save file' and 'save as' from the pull-down menus.

You can use **1- or 3-letter** amino acid code.

Space, tab and return character will be ignored.

If you use unusual character or 3-letter code, the setting job will not work properly.

If you are ready to start the setup, select '1 letter code' or '3 letter code' from the pull-down menu "Setup Project". Then the setup program will make Acs directory and give a message telling you all "Setting job has been finished.".
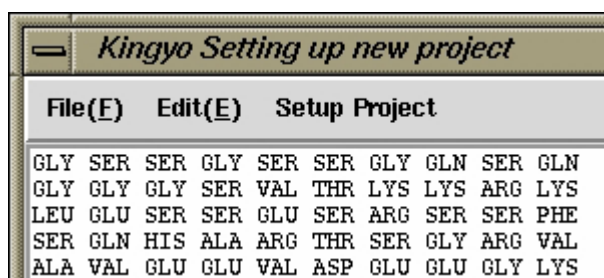
4. Instructions of startup module.
1) How to fill the spectrum list

Check box, Use or Not   Entry for file name
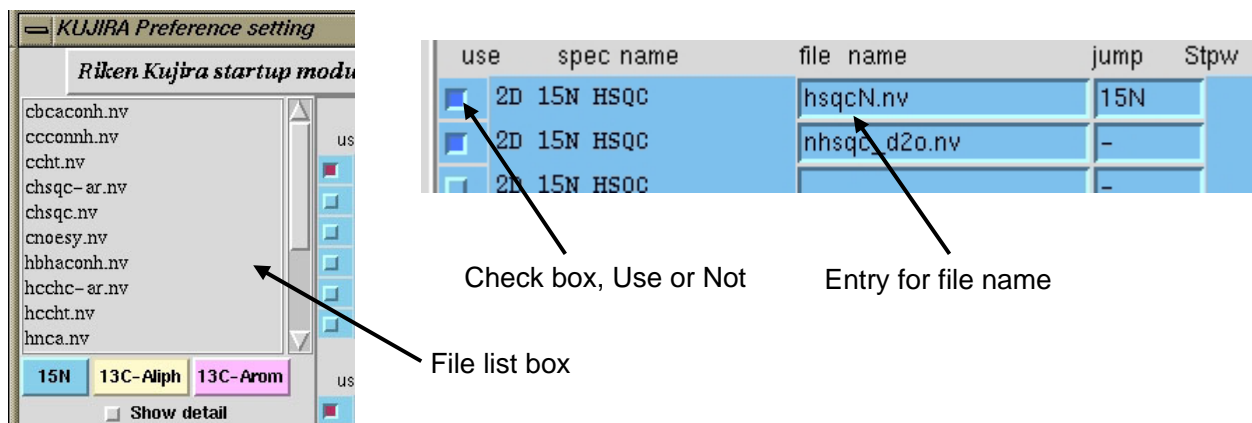
File list box

The "Startup of module of Kujira" is used for management of spectrum files that you would like to use for NMR analysis.

"File list box" shows list of spectrum all *.nv files found in the matrix directory.

Firstly you have to specify the spectrum files "*.nv" on the startup module.

Enter the file name of spectrum that you would like to use in the corresponding entry as shown above.

To use the entered spectrum for the assignment job, activate the 'check box' before start Kujira.

Kujira has standard spectrum types as follows;

spectrum types	spectrum names
2D 15N HSQC	2D ^{1}H-^{15}N HSDC
2D 13C CHSQC-all-aliph	2D ^{1}H-^{13}C HSQC for all region or aliphatic
2D 13C CHSQC-aro	2D ^{1}H-^{13}C HSQC for aromatic
HNCO	3D HNCO
HNCACO	3D HN(CA)CO
HNCOCA	3D HN(CO)CA
HNCA	3D HNCA
CBCA(CO)NH	3D CBCA(CO)NH
HNCACB	3D HNCACB
HBHA(CO)NH	3D HBHA(CBCA)(CO)NH
HNHAHB	3D HNHAHB
CC(CO)NH	3D CC(CO)NH
H(CC)(CO)NH	3D H(CC)(CO)NH
1H-15N NOESY	3D ^{1}H-^{15}N HSQC NOESY
13C NOESY aliph	3D ^{1}H-^{13}C HSQC NOESY for aliphatic
13C NOESY arom	3D ^{1}H-^{13}C HSQC NOESY for aromatic
HCCH-COSY	3D H(C)CH-COSY for aliphatic
HCCH-TOCSY	3D H(C)CH-TOCSY for aliphatic
CCH-TOCSY	3D (H)CCH-TOCSY for aliphatic
HCCH-COSY arom	3D H(C)CH-COSY for aromatic
HCCH-TOCSY arom	3D H(C)CH-TOCSY for aromatic

If you can not find the spectrum type that you want to use, you have to define the spectrum as 'Other1, Other2 and Other3'.

2) Sync-Jump function of Kujira

The function we call 'Sync-jump' is one of the most important function of Kujira. The function allows Kujira modules as described other sections, **Yellow**, **Skyblue**, **Magenta** and **Green** modules to give a command to synchronized jumping 2D spectrum strips within their sync jump attributes (Sync-Jump). The Sync-Jump achieves by the amino acid residue number, peak ID number, chemical shift, NOE peaks and atom type. This function will be very helpful to manipulate the spectrum windows without any annoying manual operations to set spectrum slice position or to calculate spectrum folding settings and other keyboard or mouse works, as well as useful to save the space of display area.

There are three attributes for the Sync-Jump function:

Attributes	*Corresponding spectrum types*
15N	^{1}H-^{15}N type
13C-al	^{1}H-^{13}C for all region or aliphatic type
13C-ar	^{1}H-^{13}C for aromatic type

For example, if you define several triple resonance spectra (such as HNCACB and ^{15}N-edited NOESY) with 15N Sync-Jump attribute, user is allowed to move their spectrum strips based on the assigned position of 2D ^{15}N-HSQC signals.

Sync-Jump based on the signal

Sync-Jump based on the signal B
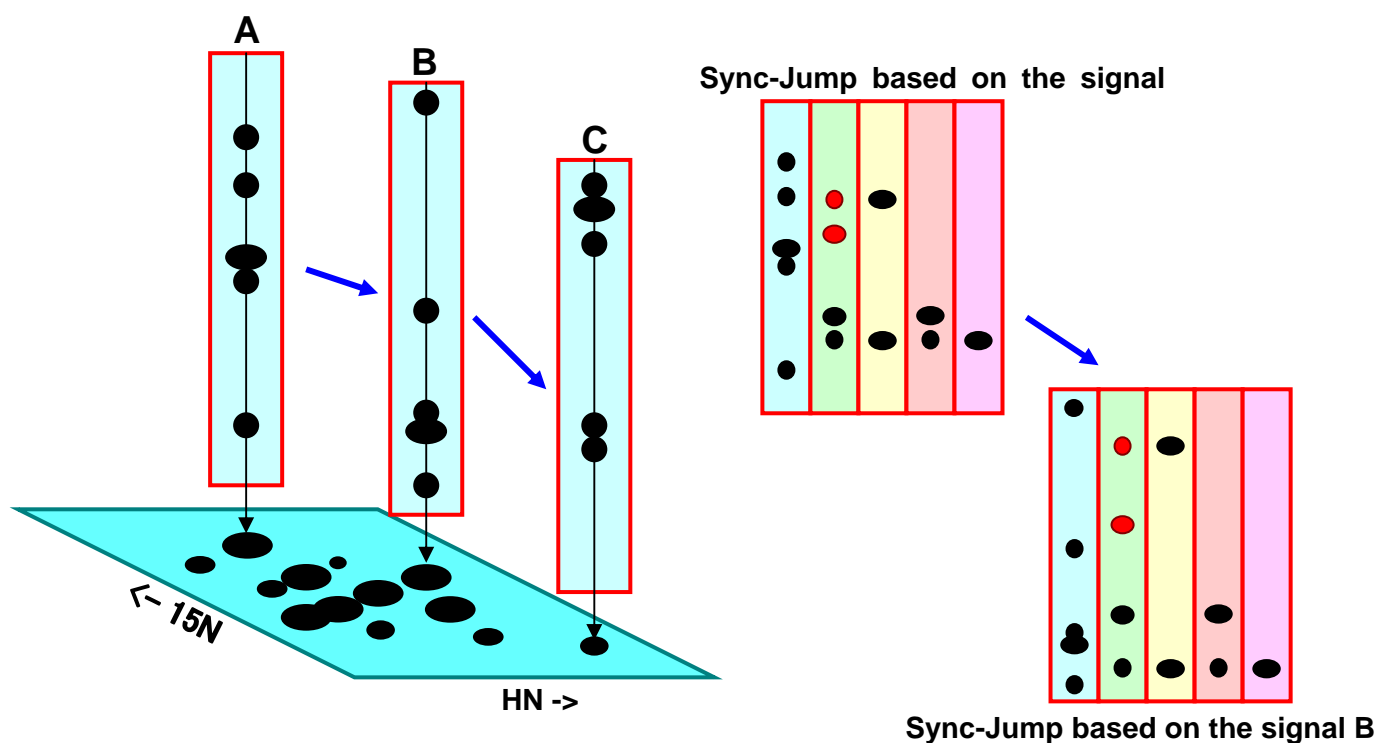
3) Settings for spectrum windows

To switch the startup module among the three sync jump mode,   there are three colored button on the left side of startup module.
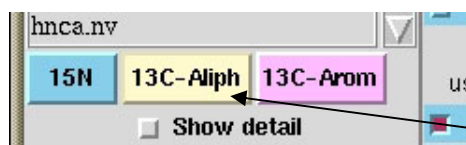
Buttons to switch sync jump modes

for ^{15}N spectra
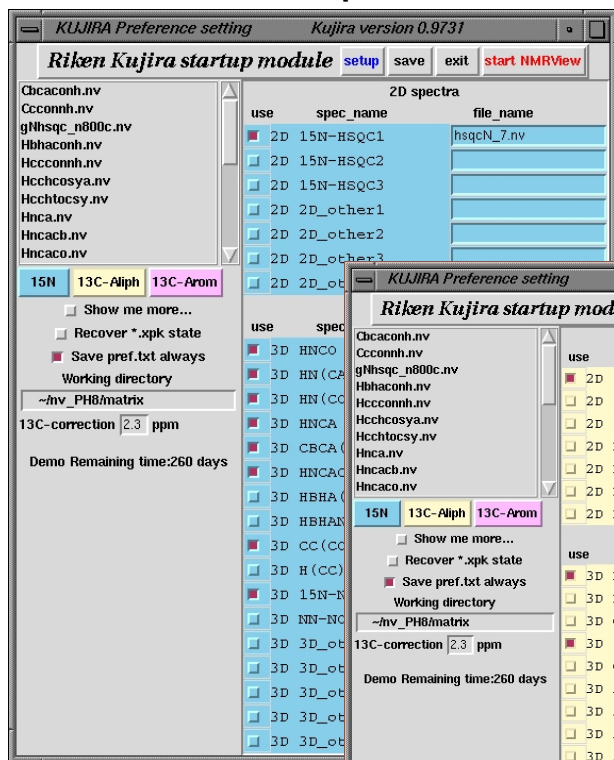
for ^{13}C-aliphatic spectra
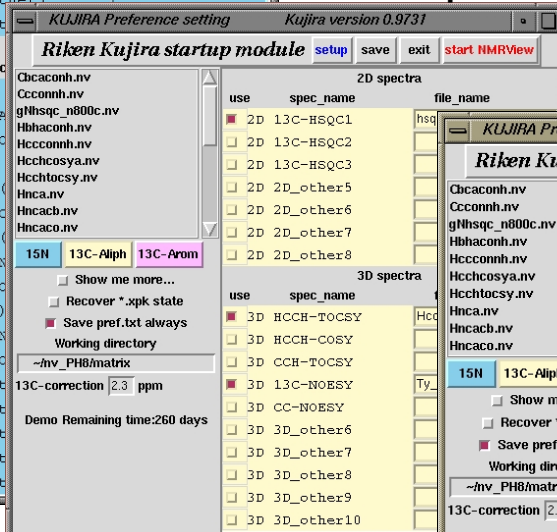
for ^{13}C-aromatic spectra
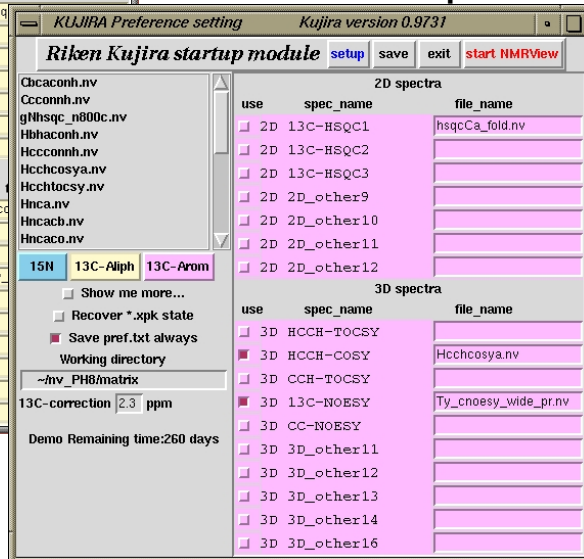
4) Spectrum window width X-axis, aspect ratio for Sync-Jump function for 2D spectra

The startup module in detailed mode shows the X-spectrum window width and aspect ratio.

1.5	1.5	pos	0/0	HN	C
1.5	1.5	pos	0/0	HN	C
1.5	1.5	pos/neg	0/0	HN	C
1.5	1.5	pos	0/0	HN	C
1.5	1.5	pos/neg	0/0	HN	C
1.5	1.5	pos	0/0	HN	H

Aspect

X-width (ppm)

If you change these values, the 2D spectra will displays the region as specified in the module by the Sync-Jump commands;
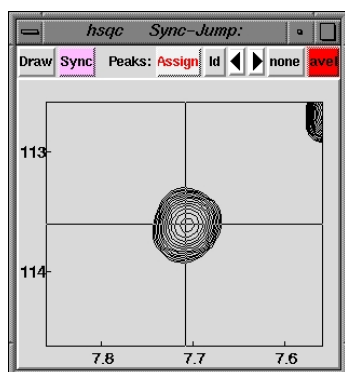
X-width: 0.5
Aspect: 80
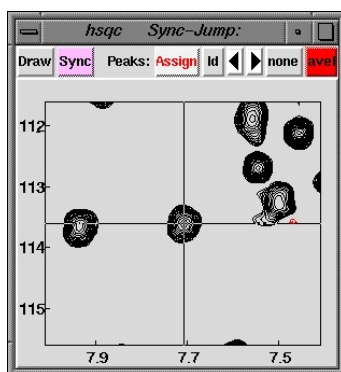
X-width: 1.0
Aspect: 80
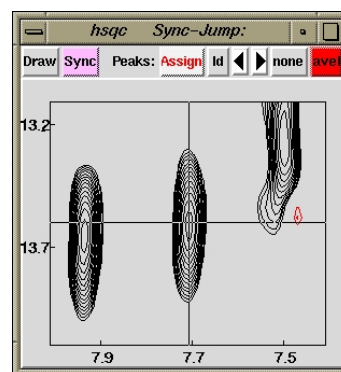
X-width: 1.0
Aspect: 20
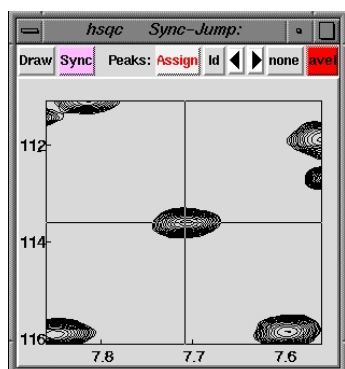
X-width: 1.0
Aspect: 200

I-7

5) Spectrum window width X-axis, aspect ratio for Sync-Jump function for 3D spectra

The values specified in the startup module can control the spectrum window width on X-axis and the aspect of X-Z axis described same as the previous section.
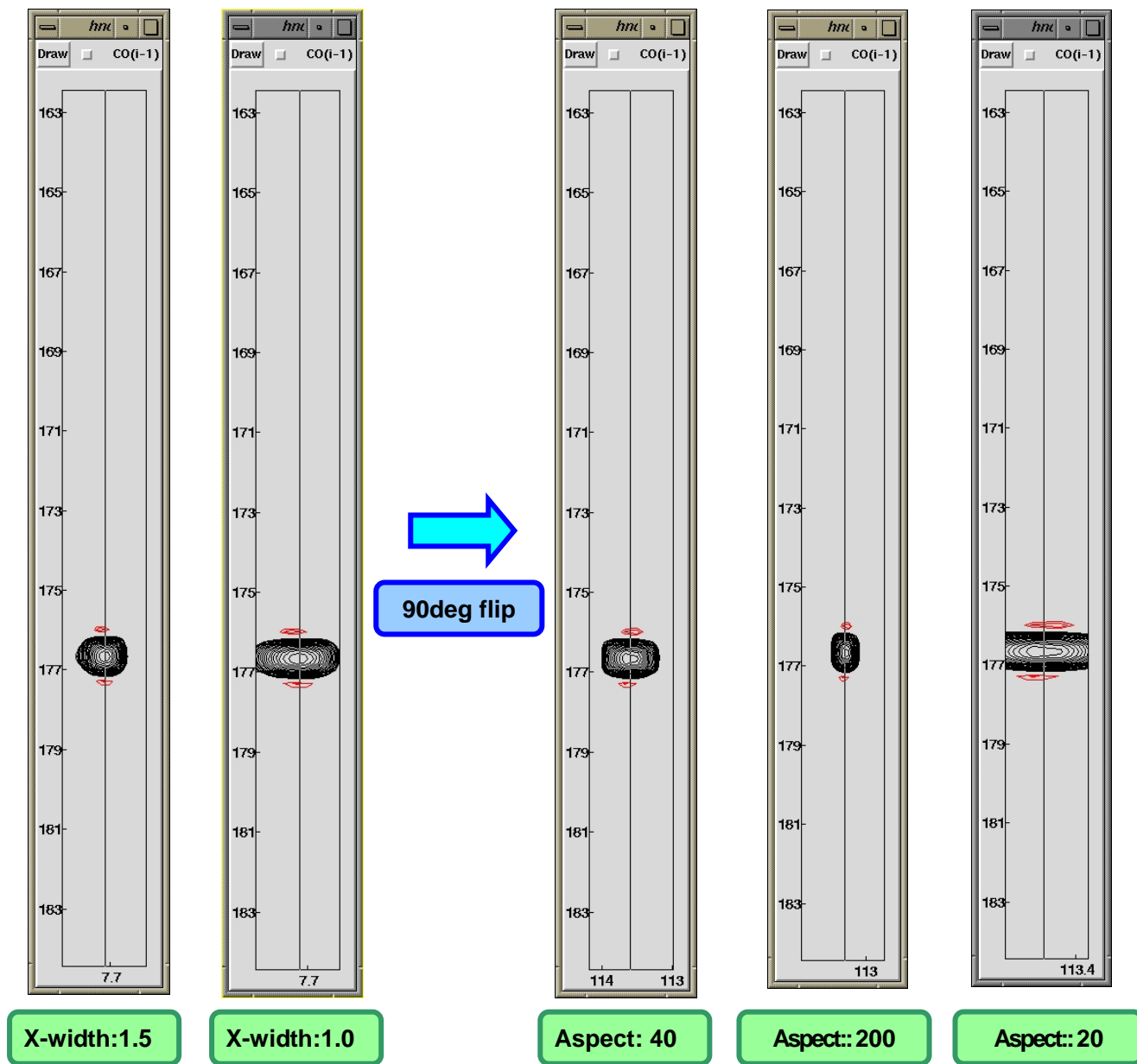
6) Store the settings to kpref.txt

If you press 'Save' button, you can save the settings to kpref.txt file in the matrix directory.

If you activate the check box 'Save setting every time', the settings are saved every after you start Kujira.

7) The other settings
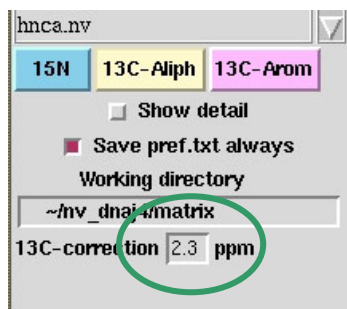
User can set correction value for ^{13}C chemical shift in the entry as shown in the left panel.   This value is not absolutely necessary for the Kujira works but it will be used for chemical shift check program and automatic assignment programs.

If user find the ^{13}C- type spectrum is slightly different from standard chemical shift, we would recommend you to adjust the offset of the spectrum before start Kujira.

5. Other functions related to "Sync-Jump

1) Switching the "Sync-Jump" action types on the spectrum windows

All the spectrum windows have two or three action types for the "Sync-Jump" function.   For 2D HSQC spectrum and 2D spectrum strip extracted from 3D spectrum have "Sync-Jump" and "Silent" modes.   The window set in the latter one can be controlled by the "Sync-Jump" commands from the responsible modules.   On the other hand, in the "Silent" mode, the window does not respond to the "Sync-Jump" command.   Only 2D HSQC type spectrum window has the "Click-and-Jump" mode.   This makes "Sync-Jump" the spectrum windows in the same sync-jump class to the desired spectrum position by directly clicking on the 2D HSQC spectrum window.
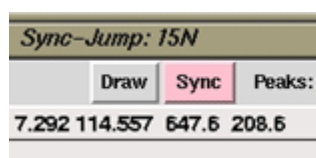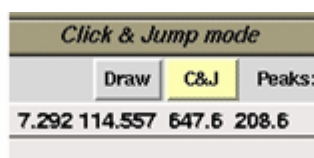
Sync-Jump   **Click-and-Jump**   **Silent**

2) Fixing the Y-axis draw region on the 2D spectrum strip extracted from 3D spectrum

There is a function to fix the y-axis draw region on the 2D spectrum strips extracted from 3D spectrum.

If you activate the check-box on the left-head of the 2D spectrum strip, then execute some sync-jump command, you will see that the y-axis draw size will be fixed with the region specified just before the execution of the command.

This function will be very useful for the NOE peak check by the CYANA result analysis module.

To reset the fixed y-axis region, deactivate the check-box, then execute the sync-jump again.   You will see the draw region has been set all region of y-axis.
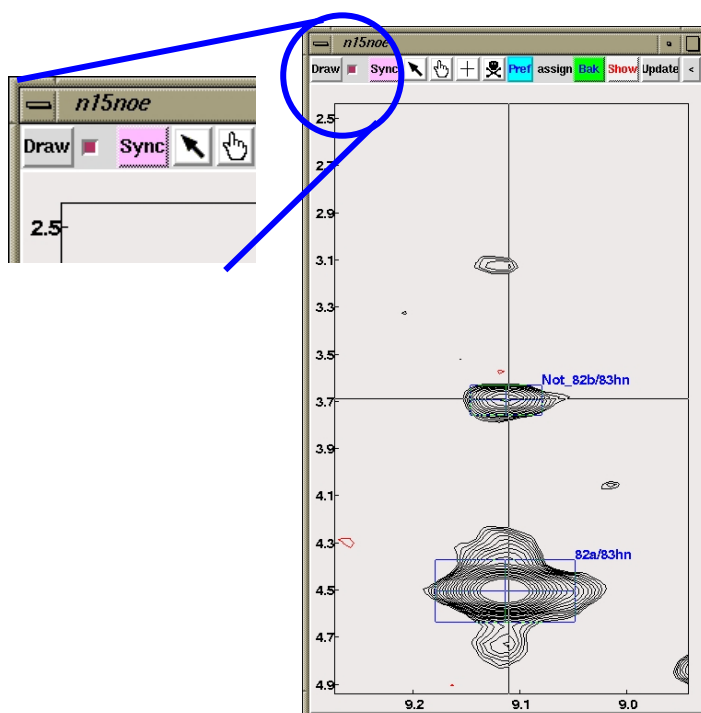
6. Directory tree of KUJIRA

/kujira

    /cshfiles

        kj1.csh   <------- csh script file to execute KUJIRA for multiple sample projects
        README.txt

    startup_kujira.tcl <------- tcl script file for starting up KUJIRA
    README.txt
    /kujira_modules <------- Tcl/Tk source files for building up the Basic GUIs

AcsManager.source:	Magenta
BBAssignManager.source:	Yellow
MainWindowManager.source:	Skyblue
StripManager.source:	Lightgreen

/plugins   <------- plugin Tcl/Tk source files and default setting files

Acs*.source :	Chemical shift table manager
BBass*.source:	Main-chain signal assignments
Cya2*.source:	CYANA result analysis
ExportCya2*.source:	CYANA input file maker
NOE*.source:	NOE peak analysis
PDB*.source:	PDB data analysis
Show*.source:	Display assignments on spectrum windows
SpecParams.txt:	Startup module default setting file
KujiraParams.txt:	KUJRA main default settings
*.def:	the default files for kpref.txt, cyana_anal.txt and cyana_input.txt
Stat_chemical_shift.txt:	Standard chemical shift table

    /startup <------- Tcl/Tk source files for KUJIRA startup

Startup*.source:	Build Startup GUI module and subroutines
Spec*.source:	For making Spectrum windows

    /bin   C-program binaries
    /images   picture files

7. Main-chain assignment

1) Requirements

Spectrum	target signal	axis order	axis labels
*3D HNCO	CO(i-1)	^{1}H(acq)-^{13}C-^{15}N	HN-C-N
3D HN(CA)CO	CO(i)	^{1}H(acq)-^{13}C-^{15}N	HN-C-N
3D HN(CO)CA	Cα(i-1)	^{1}H(acq)-^{13}C-^{15}N	HN-C-N
3D HNCA	Cα(i)	^{1}H(acq)-^{13}C-^{15}N	HN-C-N
*3D CBCA(CO)NH	Cβ(i-1)	^{1}H(acq)-^{13}C-^{15}N	HN-C-N
*3D HNCACB	Cβ(i)	^{1}H(acq)-^{13}C-^{15}N	HN-C-N

The spectra indicated with asterisks are essential for the assignment jobs.

If you do not have HNCA and HN(CO)CA, define HNCACB and CBCACONH spectra as Cα(i) and Cα(i-1) type spectra instead of them.

Caution: Be careful to set axis order and axis labels, otherwise the Sync-Jump jobs will not work properly.   Do not use same axis labels in a spectrum, for example, H-H-C.   This will fail to pick and save the peak information properly.

2) making peak table, assign_BB.txt

Before start the mainchain signal assignment, user has to make a peak id table file "assign_BB.txt".

User can use 2D-HSQC, HNCO or HNCOCA to make the peak ID table.

Firstly, user has to make *.xpk file.

Press "Setup" button on the yellow module, then you will see a small window as shown below.   Select one of the xpk file types, then press "Setup" button to start the making job.

If you don't see any error message, the job has been successfully finished.

Main-chain signal Assignment
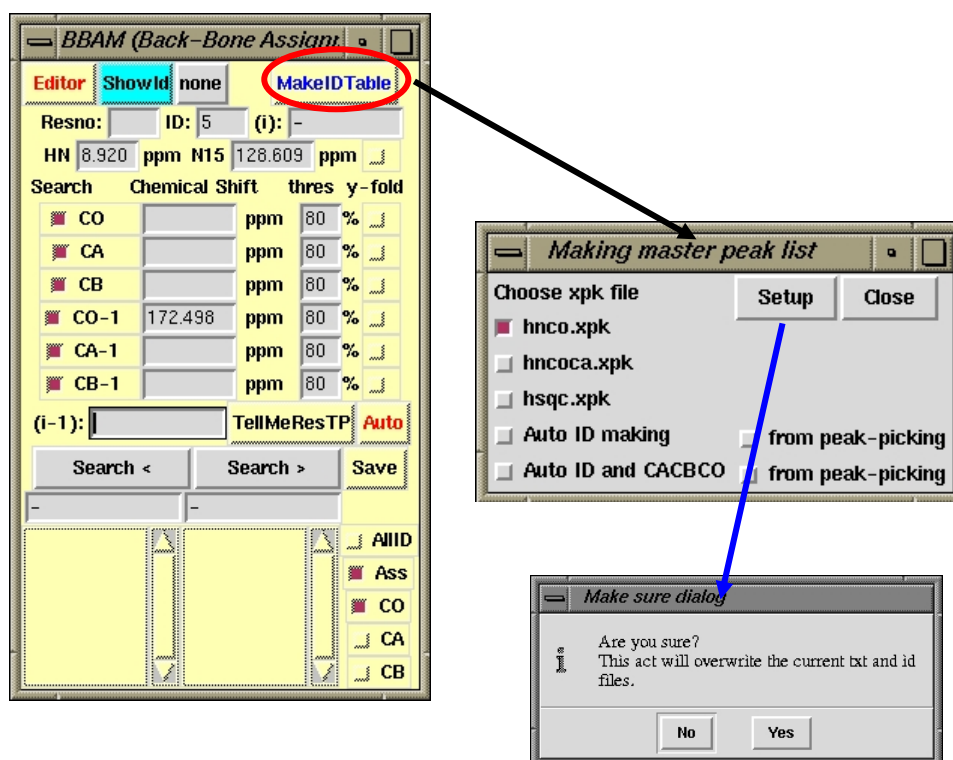
3) Edit master peak table, assign_BB.txt

If you successfully made assign_BB.txt, you can find the file in /matrix.

If you press "Editor" button on the yellow module, you will see a text editor as shown below.

Kujira gives a command to control 2D and 3D spectrum windows based on the peak ID and chemical shift described in the assign_bb.txt file. Using the above editor, user can add and delete Peak Ids just like a text editor.

Caution: Do not duplicate ID number. Use only integer for ID number.
Make sure the number of columns are the same on each lines.

User can use 2D ^{1}H-^{15}N HSQC for the deletion and addition of Peak ID.
If you press "Id" button on the header of HSQC window, the peak boxes will appear.
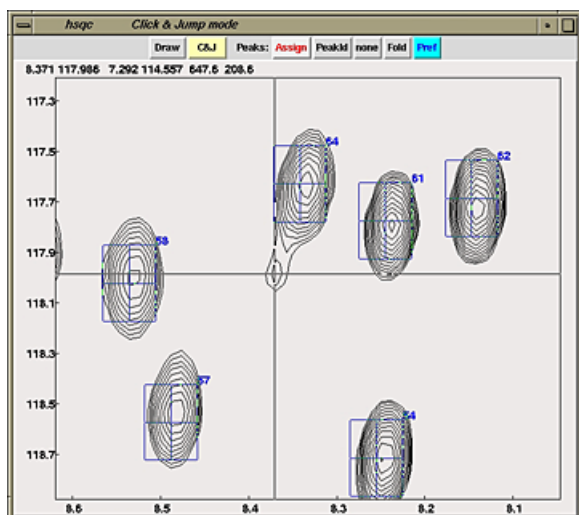
Then if you change the cursor state to "peak delete", "peak add" or "peak adjust", you can add and delete peak boxes.

If you press "Save" button, Kujira will store the current peak Ids on the HSQC window into assign_bb.txt file.

Main-chain signal Assignment

4) Test the sync jump and offset adjustment.

Align the spectrum windows as shown the left panel. Press increasing and decreasing button on the
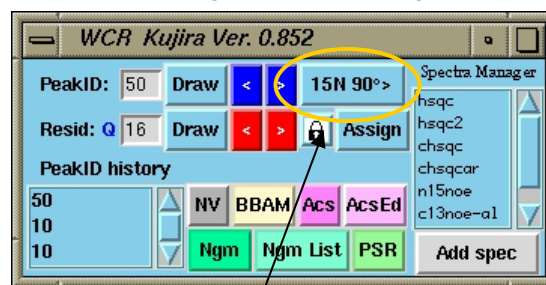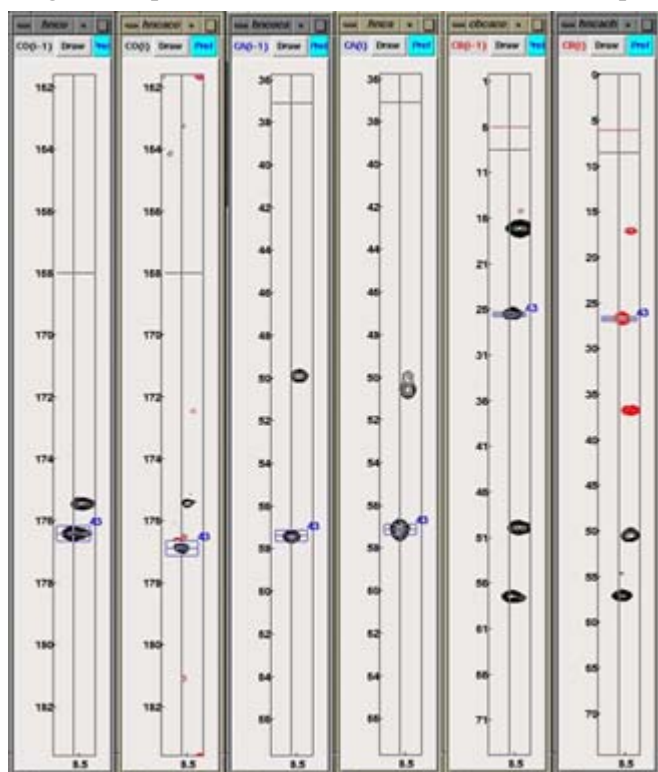
¹⁵N 90deg flip

skyblue module, and check the spectrum strips are moving depending on the peak positions as described in the assign_bb.txt file.

Press **"¹⁵N 90°>>"** button, then check the spectrum strips are flipping 90 degree with y-axis.

Then check the black and red cross-hair lines are moving on the other spectrum strips. If you can not see the cross-hair moving, confirm the spectrum labels are same.

Check the all signals are exactly on the center of strip either on x- or z-axis. If you look at one spectrum and find all of the signals are slightly out of the center, try to change the x- and z-offset values. By clicking "Pref" button, you will see a small window "Preference". Enter new offset value and press "Draw" button on the skyblue module to check whether the signal is on the center of strip.

Main-chain signal Assignment
5) ¹³C signal identification for main-chain signal assignments

The next step for the main-chain signal assignment is to identify ¹³C signals for each peak ID. In Kujira, user is allowed to directly detect the chemical shifts of the carbon signals, Cα(i-1), Cα(i), Cβ(i-1), Cβ(i), CO(i-1) and CO(i) by clicking on the 2D spectrum strips. Header label of each strip shows which ¹³C signal type is available for the detection as shown below.
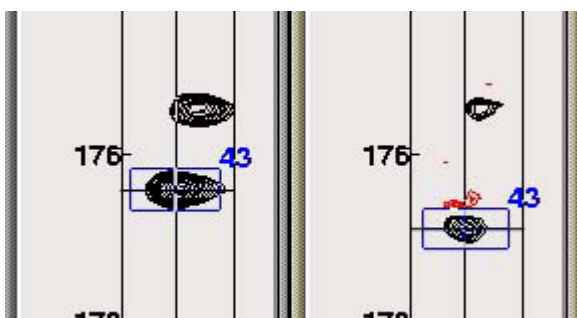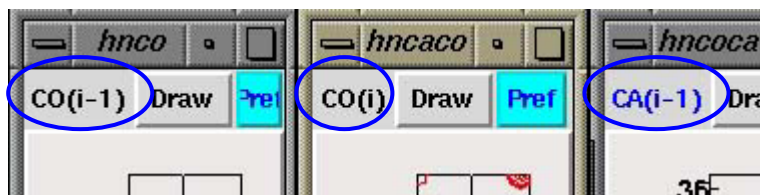
If you align all spectrum strips and show a certain peak ID, you will be able to input the chemical shift value to the yellow module by middle-button-clicking on the center of target signal.
The detected values will be immediately input the yellow module as the left-below panel.

By pressing "Id" button on the head of the yellow module, user can check exactly where you have specified on the spectrum as blue boxes shown on the above-left panel.

Important: To save the values in the "assign_bb.txt", user has to press "Save data" button.

If you press "Tell me res-type" button, Kujira predicts amino acid type for (i) and (i-1) residues. The predicted residue type will appear on "Res-type" entry as one-letter code. The higher probability of predicted amino acid type, the more left side of the entry.

If you find the peak is obviously noise, artifact or minor conformation, you can define it in the "Type(i)" entry. For example, if the signal seems to be noise, type "noise" in the "Type(i)" entry so that the automatic assignment program will ignore the signal.

Main-chain signal Assignment

6) Sequential assignment

The chemical shift values that you have input so far are saved in the text file, matrix/assign_bb.txt. Kujira periodically takes backup file in matrix/backup directory at every ten actions of edit.

Press "Editor" button on the yellow module to open the "Kujira Assignment Table Manager". The editor displays the entered chemical shift values where the columns are corresponding to "Residue number", "Residue type", "Peak ID", "HN", "15N", "CO(i)", "CO(i-1)", "CA(i)", "CA(i-1)", "CB(i)" and "CB(i-1)". User can sort the columns by pressing the buttons just above them.

Kujira Assignment Table Manager:/home/naohiro/nv_ubiq/matrix/assign_BB.txt

File(E) Edit(E) Tools Check Assign Show Segments   Total: 79 assigned: 72 not_assigned: 7 gomi: 3   Sequence

Quick Save Undo Redo Buffno: 0 (Max:0) Sort: increasing decreasing   Close

| #Peak	Rank	ResNo	Score	Restp(i-1)	Restp(i)	HN	N	CO(i-1)	CO(i)	CA(i-1)	CA(i)	CB(i-1)	CB(i)	conn(i-1)	conn(i+1)
0	0	999	1.000	-	gomi	8.308	126.500	178.305	999.990	999.990	999.990	999.990	999.990	-	-
1	0	77	1.000	LDC	V	9.078	126.237	173.053	171.748	51.420	58.382	41.744	32.500	13	20
2	0	57	1.000	QHERWM	DLNFY	8.465	125.396	173.210	174.250	53.556	51.740	26.749	39.142	17	19
3	0	22	1.000	T	LC	8.735	125.057	171.332	172.576	59.904	50.508	67.146	44.003	28	27
4	0	82	1.000	REHWMCQ	G	8.400	124.553	174.474	171.244	54.170	42.908	28.255	999.990	29	67
5	0	52	1.000	LIYFD	FLYCD	8.738	124.455	173.339	172.273	56.598	54.137	38.731	41.264	25-78	49
6	0	69	1.000	IY	Q	7.571	124.441	172.079	173.423	60.197	51.192	34.212	29.214	51	43
7	0	65	1.000	S	D	7.870	124.307	175.943	175.120	58.822	55.027	60.299	38.046	76	64
8	0	999	1.000	-	gomi	7.604	124.351	177.251	999.990	999.990	999.990	999.990	999.990	-	-
9	0	80	1.000	RWHMEKQC	L	8.284	124.267	172.983	175.037	53.409	52.613	28.940	40.032	14	29
10	0	50	1.000	RWHMEKQC	LDC	8.711	124.108	171.435	172.878	52.824	50.919	29.351	43.455	23-18	25-27
11	0	7	1.000	S	G	8.221	123.801	172.498	170.881	56.247	43.182	61.805	999.990	53-77	44-63
12	0	28	1.000	S	DYFL	7.957	123.586	172.237	173.968	55.018	53.794	60.983	38.457	58	22-21
13	0	76	1.000	KMWHR	L	8.179	123.374	171.599	173.060	54.053	51.192	30.241	41.881	79-23-10-30	1
14	0	79	1.000	LDN	RHEMKQWC	8.501	123.318	175.497	172.939	51.683	53.246	40.237	28.940	20	9-41-10
15	0	38	1.000	I	QHEWC	8.451	123.194	175.795	176.490	63.766	57.560	34.349	25.380	37	45
16	0	35	1.000	MVK	A	7.874	123.135	178.102	177.841	56.832	52.973	31.131	15.246	50-42	63
17	0	56	1.000	KMREH	QHERWC	8.565	122.854	172.223	173.221	52.268	53.657	31.953	26.680	30	2
18	0	49	1.000	RWHMKEC	HMQERWK	8.392	122.837	173.683	171.446	54.375	52.630	29.145	29.351	54	10-5-78
19	0	58	1.000	DLYFN	-	8.341	122.786	174.238	173.140	51.859	53.643	39.005	29.556	2	41
20	0	78	1.000	VMI	L	8.052	122.734	171.734	175.481	58.178	51.466	32.432	40.306	1	14
21	0	17	1.000	T	G	7.741	122.768	173.118	171.587	59.056	43.045	66.940	999.990	46-22	40-63-31-30
22	0	29	1.000	DYFL	T	7.752	122.644	173.982	174.331	53.731	57.286	38.526	68.720	12-2-21	36-35
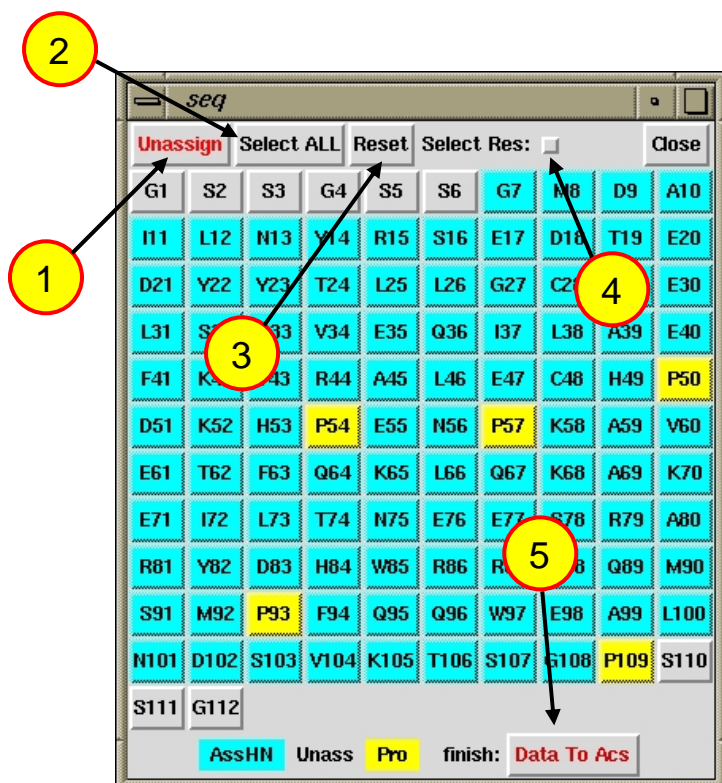
If you press "Sequence" button, you will see sequence board window. Functions on the window are;

1: Unassign selected residues.   2: Select all residues   3: Reset selection of residues
4: Checkbox for enabling to unassign selected residues.   5: Data exportation to Acs (magenta) module.

Main-chain signal Assignment
The right panel shows the selected residues, 104-108, are indicated by dark gray
If you press "**Quick Assign**", you will see an window as shown below.
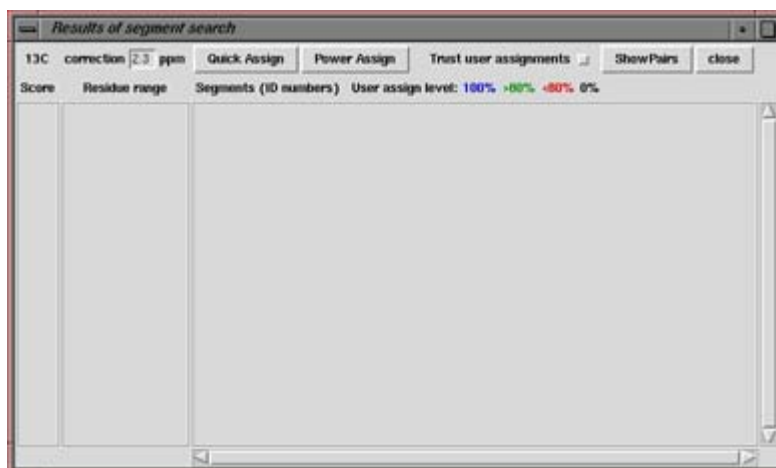
In the window, user can define the chemical shift tolerance for CO, CA and CB in ppm which will be used for the automatic sequential assignment program "QuickAssign". The default values are 0.3 ppm respectively. The systematic chemical shift correction is also applied to all carbon chemical shift in the "13C correction" entry. If you use Xwin-NMR (Bruker) or JEOL for the offset calibration, you probably have to set the value around 2.3.

If you press "**QuickAssign**" button, the automatic assign program will run. The calculation time is within a few second (depending on the number of signals).
Caution: If the peak table contains too many minor signals and artifacts, program will crush.

If successfully the program runs, you will see:

The left column is penalty score of the found segment. The middle one is predicted begin-end residue number. The right one shows peak-ID connectivity of found segment.
The penalty less than 10.0 are shown but better result will be less than 5.0.
Important: If more than two possible segments are found, the residue-range will be colored by red. This is mainly because the peak ID list contains minor signal showing the same chemical shift for CO, CA or CB. If you carefully check the peak pattern then you find it a minor signal, you should define the signal as "minor" in the yellow module so that you will be able to avoid the multiple assignment

Main-chain signal Assignment

Segment detail

Very good | Penalty:1.531 | Assign | Unassign | Close

Peak ID	Rno	Rtp	User	dCO	dCA	dCB	ResidueType(i-1)	ResidueType(i)
---	60	G	---	---	---	---		
48	61	R	61	---	---	---	G	R
24	62	T	62	-0.00	-0.03	0.07	MREKQH	T
56	63	L	63	-0.01	-0.02	0.21	T	L
76	64	S	64	0.02	0.04	0.00	LFDIYC	S
7	65	D	65	-0.00	0.10	-0.07	S	D
64	66	Y	66	0.01	0.02	-0.14	DLYFI	Y
62	67	N	67	-0.01	0.14	-0.20	FYIDL	N
51	68	I	68	0.02	-0.01	-0.14	NMCKYF	I
6	69	Q	69	0.05	-0.10	-0.27	IY	Q
43	70	K	70	-0.00	0.02	-0.07	QHRMWKE	K
69	71	E	71	0.03	-0.07	0.00	KMCWV	E
68	72	S	72	0.00	0.22	-0.18	QWC	S
57	73	T	73	0.02	-0.12	-0.27	S	T
71	74	L	74	0.01	-0.27	0.00	T	L
79	75	H	75	-0.03	0.28	-0.07	LDC	H
13	76	L	76	0.02	0.01	0.02	KMWHR	L
1	77	V	77	-0.01	0.23	-0.14	LDC	V
20	78	L	78	-0.01	-0.20	-0.07	VMI	L
14	79	R	79	0.02	0.22	-0.07	LDN	RHEMKQWC

result.
If you double-click one of the segments, another small window will appear. The columns are corresponding to "Peak ID", "Residue number", "Residue type", "delta CO", "delta CA", "delta CB", "Predicted residue type(i-1)" and "Predicted residue type(i)". If you don't find any problem, press "Assign" button to determine the assignment result.

The determined segment will turn blue.

Results of segment search

13C | correction 2.3 ppm | Quick Assign | tol: CO 0.3 | CA 0.3 | CB 0.3 | Fix assigned residues: ☐ | ShowPairs | close
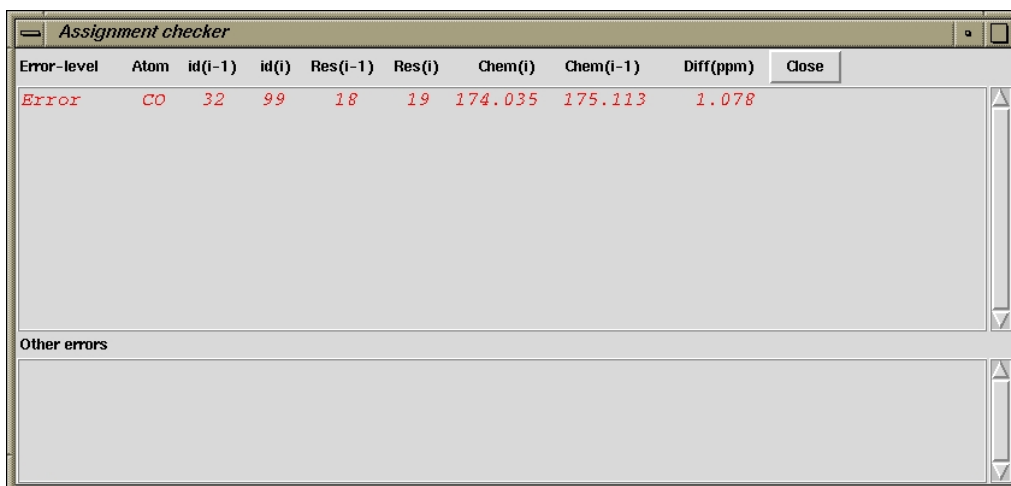
Penalty | Residue range | Segments (ID numbers) User assign vs result: 100% >80% <80% 0%

Penalty	Residue range	Segments (ID numbers)
1.124	7-17	11-44-32-74-52-34-72-66-35-46-21
2.576	19-25	39-73-28-3-27-60-55
2.112	27-30	58-12-22-36
2.392	32-43	33-26-50-16-42-37-15-45-65-70-23-40
2.104	46-54	75-59-54-18-10-25-5-49-61
0.869	55-59	30-17-2-19-41
1.531	61-82	48-24-56-76-7-64-62-51-6-43-69-68-57-71-79-13-1-20-14-9-29-4

Important: If the segment length is less than 4, you should confirm the assignments by carefully checking residue type (i-1) using CC(CO)NH spectrum before the determination.

If you activate "Fix assigned residues:" check box in the red circle, you can fix the assigned residues in the table during the QuickAssign calculation. This function will help the QuickAssign works to find another possible segment by excluding the assigned peak ID.

Main-chain signal Assignment
If you press "CheckAssign" button on the header of the editor, you will find another window.

If the module finds very serious problem it will give an error message, for example, a pair of assigned residues give a large differe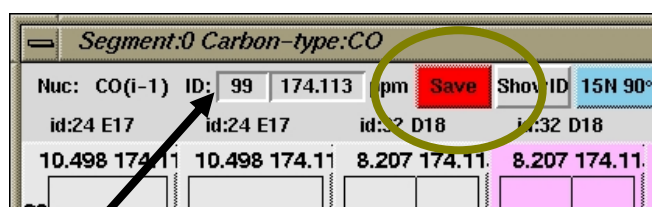nce between intra-residual and sequential shift values more than 0.6 ppm. On the other hand if the module finds noticeable problem, it will give a warning, for example, the amino acid type of assigned residue does not match with the type predicted from the assigned chemical shift values.

Important: Ideally to say in the final sate f assignment, there should be no error nor warning.

If you double-click one of the error or warning messages, another window will appear as shown left below panel.
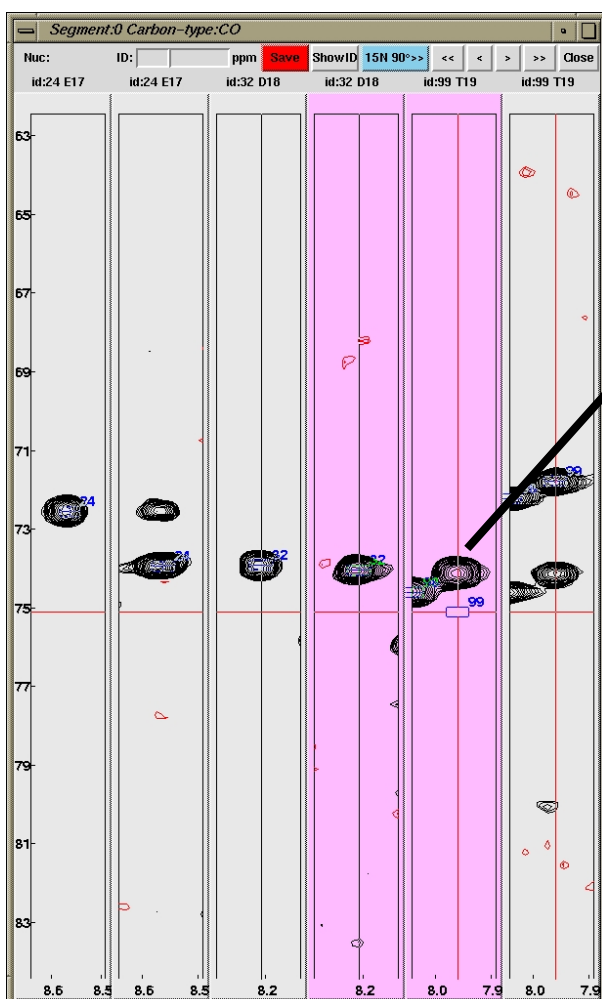
mouse middle button

The sequential (i-1) and intra-residual (i) spectrum strips are aligned side by side. Assigned residue number and peak ID are indicated on the header of the spectrum strips. The sequential connectivity with problem found by the CheckAssign are colored by magenta. User is allowed to correct shift value on the module.   If you place cross-hair on the signal and click mouse-middle-button, y-axis value will enter in the entrance on the header of the window.   To store the shift data to assign_bb.txt, press "Save" button on the module.

Main-chain signal Assignment
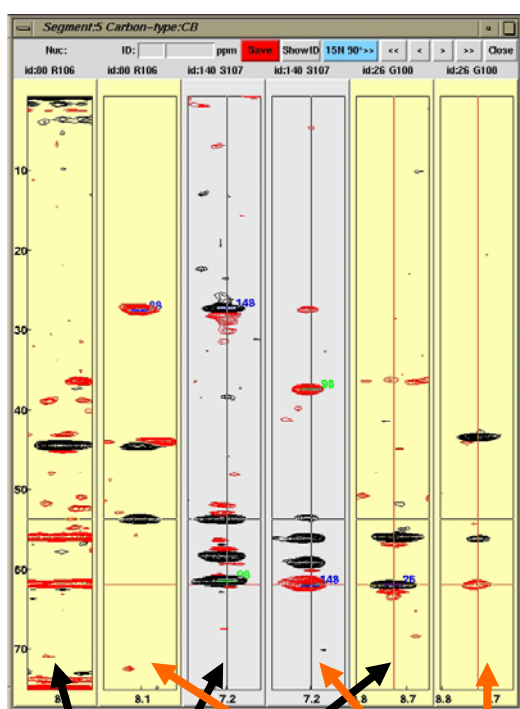## 7. Confirmation of sequential connectivities

To confirm sequential connectivity of peak IDs that you have assigned, user is allowed to show sequential (i-1) and intra-residual (i) spectrum strips aligned side-by-side. If you press "Show Segment" button on the Assignment editor, a small window will appear as shown below.
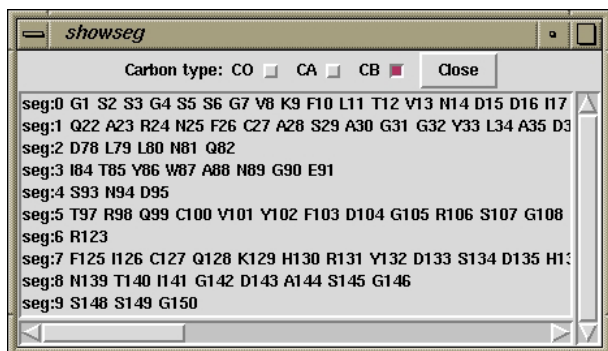
The segments shown in the list box indicate sequentially aligned amino-acid types and residue numbers expected from sample sequence. Before checking the sequential connectivity, user has to select one type of nucleus among CO, CA and CB using check-boxes on the head of window. If you double-click one of the segment, another large window will open as shown below.

For instance, if you select "CB", you will see six spectrum strips, where 1st, 3rd and 5th strips are corresponding to CBCA(CO)NH while 2nd, 4th and 6th ones are HNCACB.
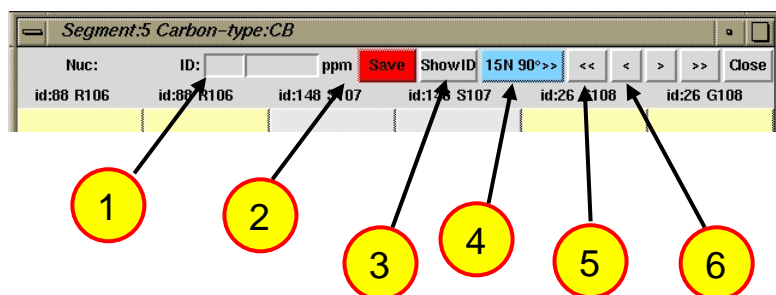
Buttons and their functions of the windows are:
1. Entry for corrected shift value. **2.** Save button. User has to press this button if user would like to store corrected shift value to assign_bb.txt file. **3.** "ShowID" button. Display stored shift position on the strip with blue boxes. **4.** 90 degree flip button. Used for flipping spectrum strips 90 degree at x-axis. **5.** Jump button to 5 residues backward. **6.** Jump button to 1 residue backward

CBCA(CO)NH   **HNCACB**

If you find that stored shift position is wrong, you can correct the value manually. Firstly place cross-hair right on the peak, then press mouse-middle button. You will see shift value on y-axis are entered in the entry of window header. To store the corrected shift value, press "Save" button.

II-5
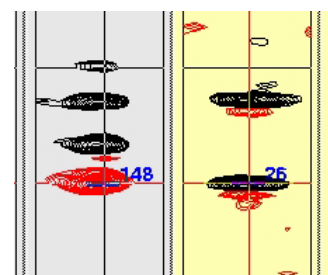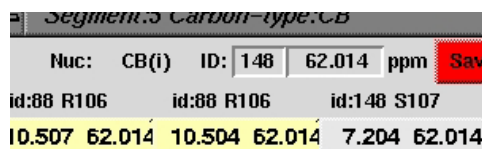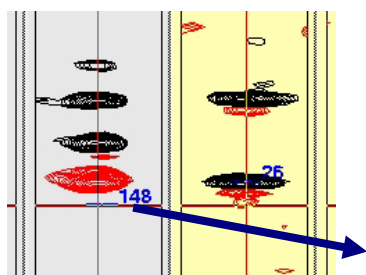
Main-chain signal Assignment

7) Function to automatically identify CA, CB and CO signals on each peak ID

User also allows to execute automated CA/CB/CO signal identification for one peak ID by pressing **"Auto"** button on the yellow module.

Main-chain signal Assignment
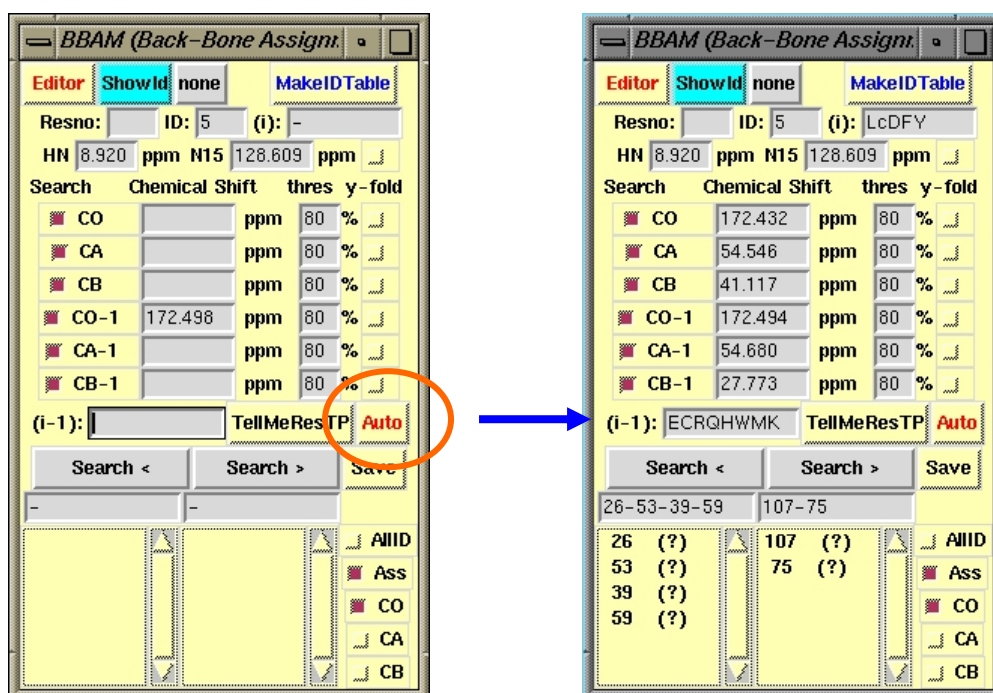
8) Exportation of assigned data to Acs (magenta module)

In Kujira, the main-chain assignment modules and chemical shift table module are working nearly independently (this will be integrated soon). Therefore user has to transport data in the assign_bb.txt to Acs directory.

Acs directory contains the most important information of Kujira, accommodating sequence data and chemical shift data. If you go matrix/Acs, you will find the files, acs.*. The chemical shift of each residue is stored in the acs files. the extension number is corresponding to the residue number.

To export the data in the assign_bb.txt file, press "**Data to Acs**" button on the sequence board, and follow the coming messages.

Caution: This act will overwrite the Acs files with new ones, so please take backup of the directory before do this job.

If you don't see any error message the data exportation was finished.

You can confirm the jobs worked properly if you press "**Jmp**" button on the magenta module, or "**Assign**" button on the head of Hsqc window.

The magenta module in the upper-left panel shows the assignment data of residue Thr80. The shaded characters indicate average value of chemical shift data found in the BMRD database.
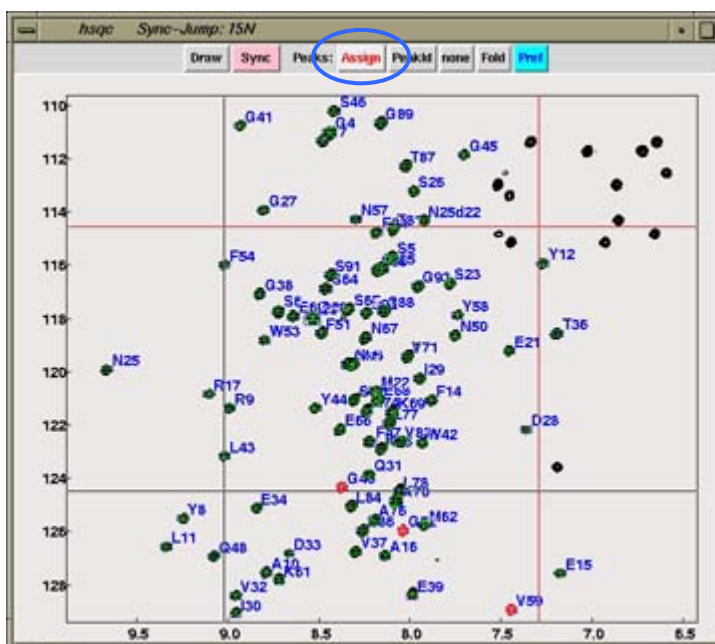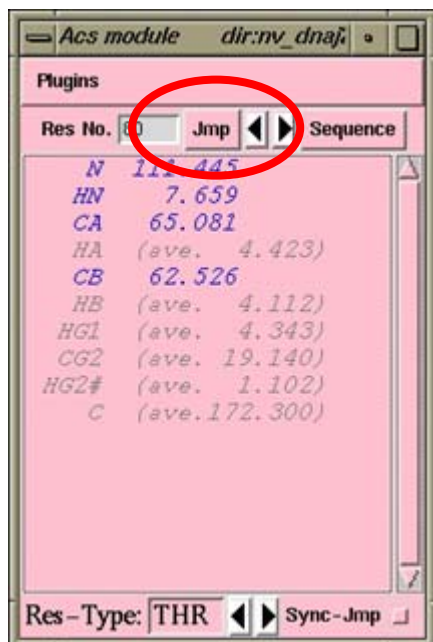
Main-chain signal Assignment
9) Automated Hαβ signal assignments using ¹H-¹³C HSQC and HBHA(CO)NH

If you have completed the sequence specific main-chain signal assignments, now you can run the program "AutoAssignHBHA". The program can export the HN, ¹⁵N, Cβ, Cα and CO that you have assigned into the Acs table as described in the previous chapter, as well as automatically assign the Hα and Hβ signals using 2D ¹H-¹³C HSQC (involving aliphatic region) and 3D-HBHA(CO)NH spectra.

The program will pick peaks of 2D ¹H-¹³C HSQC and 3D-HBHA(CO)NH spectra, then try to find the most probable signals as Hα and Hβ using the amino acid type information and the chemical shifts for HN, ¹⁵N, Cβ and Cα that you have assigned. The assignment completeness and accuracy
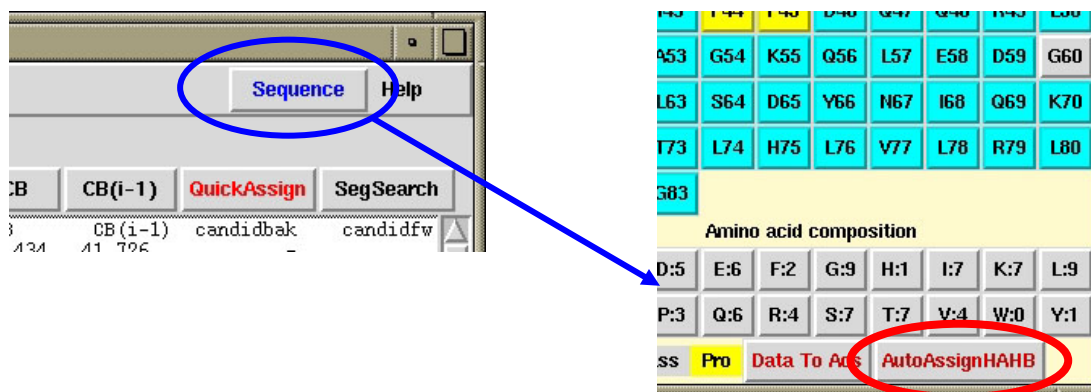
are ~90% and ~80% depending on the quality of the spectra.

To run the program, Just follow the steps:

1. Firstly, open the 2 D-¹H-¹³C HSQC、 HBHA(CO)NH spectra, then optimize their threshold.
2. Open the sequence window for main-chain assignment by pressing the yellow "Sequence" button.
3. You will be asked if you are sure, then if you press "yes" , a terminal window will open and runs for a few ten seconds.

If you would like to confirm the assignments, press "Assign" button on the 2D ¹H-¹³C HSQC window.
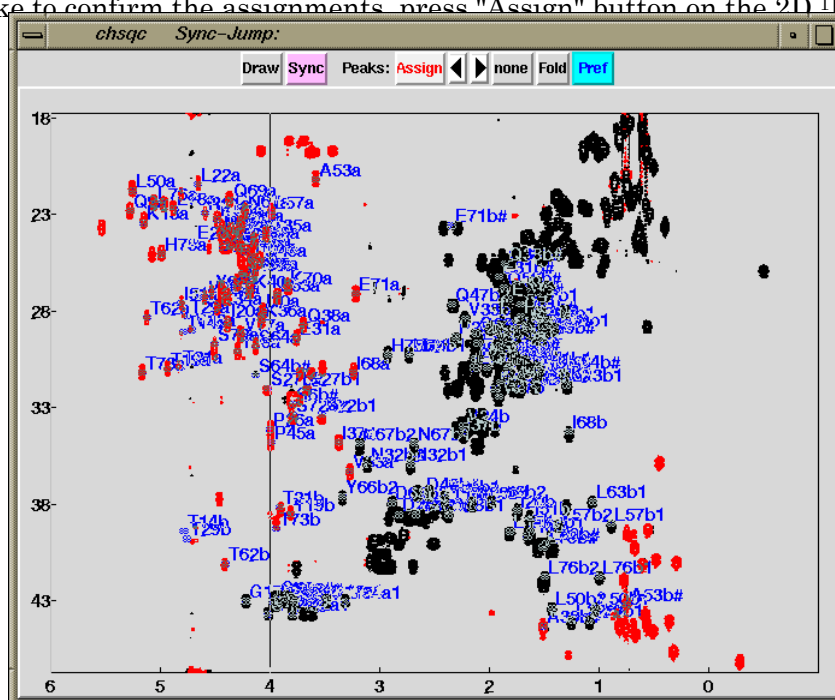
8. Acs module for managing chemical shifts and chemical structure specific assignments
1) How to use magenta (Acs) module.
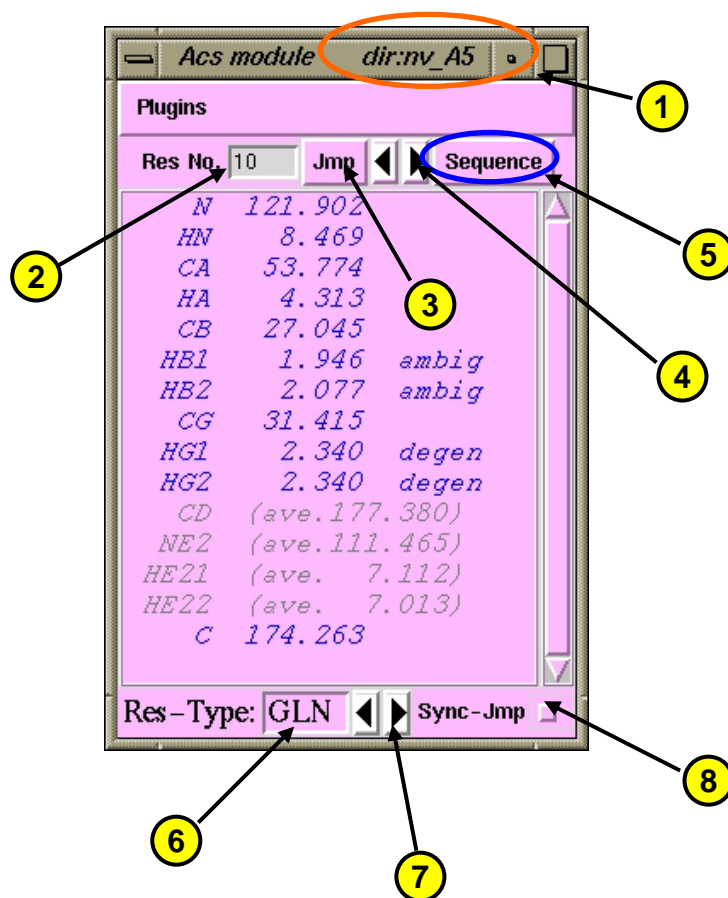
Acs (Amino acid specific Chemical Shift table, magenta)　is one of the most important module of Kujira.　The module has a lot of functions, for storing chemical shift data of all signals, controlling spectrum strips based on the shift data, comparison of shift data to standard values, and exportation to a text file with another format.

1: displaying the job directory
>   where you store nv files and Acs directory.

2:　entrance of residue number
>   If you type the residue number and press return key or press "jmp" button, the module will display all shift data of the residue.

3:　jump button
>   move to the residue number

4:　buttons for increment and decrement of residue number.

5:　button to call for sequence board
>   If you press this button, "Sequence board" window will open as shown in the next page.　If both HN and ^{15}N shift values are assigned, the button of the residue will be painted by cyan. Pro residue is always yellow.　If you press the button on the sequence board, you can move the magenta module to the desired residue number.

6: entrance for amino acid type
>   If you type one or three code of amino acid type and press buttons 7, the module will display all shift data of the nearest residue number.

7: buttons for moving to another amino acid type

8: check box for enabling and disabling "Sync-jump mode" of the module
>   If you activate the check box, the module gives sync jump command to ^{15}N type spectra. The spectrum strips in the 15N Sync jump mode will synchronizedly jump as changing residue number and residue type of the magenta module.
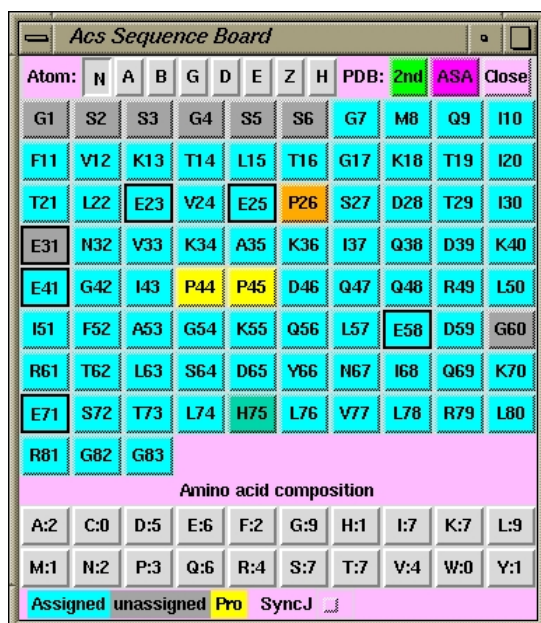
Sequence board window of magenta module

If you have not assigned any shift value of the residue, you will see the grayed characters showing atom types and their standard chemical shifts which are derived from BMRB database.

2) How to enter shift value to the magenta module

If you double-click one of the atom type, you will see a small window named "Acs Editor".
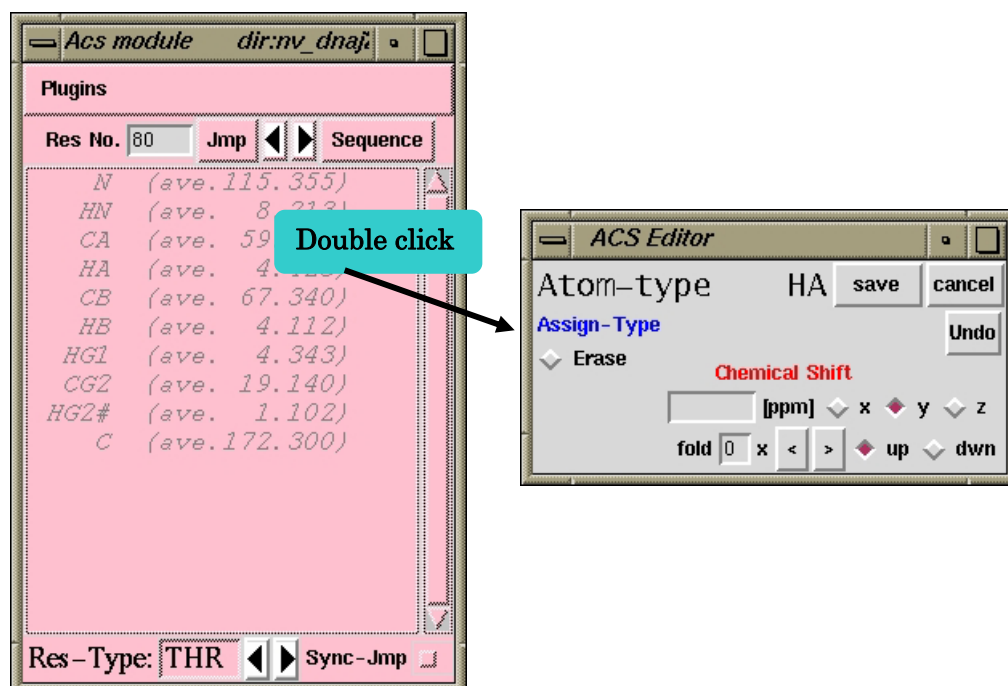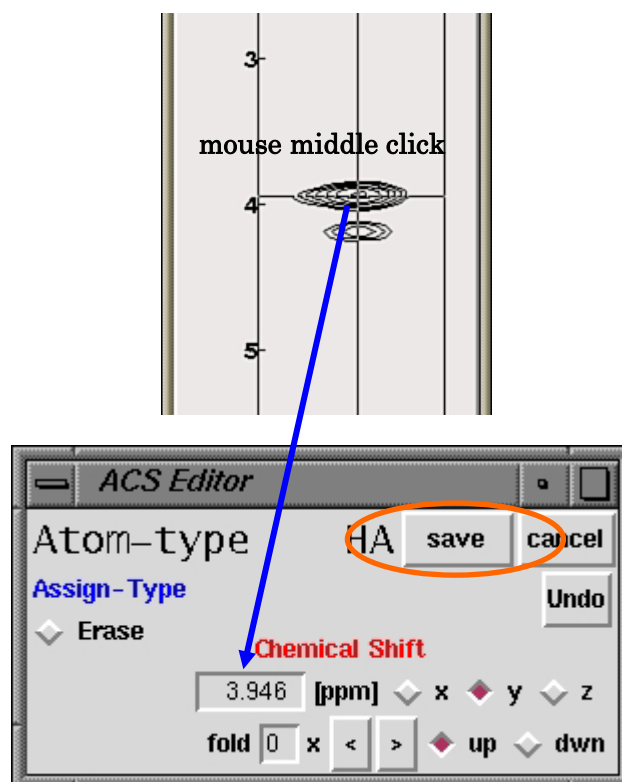
mouse middle click

If you place the cross-hair on the spectrum strip where you would like to get chemical shift, press mouse middle button.   You will see the shift value enter on the entrance of the editor.   If you want to enter value on x- or z- axis rather than the default y-axis, activate check box, x or z on the editor. **Press "Save" button to store the entered shift value.**   User is allowed to take the entered value back once by pressing "Undo" button.

If you store the shift data, you will see the atom data change blue. The magenta module always monitors the difference between the entered shift and the standard one.

If the difference is larger than twice or three times of standard deviation, the shift data turns to orange (warning) or red (error), respectively.

3) To calculate the actual chemical shift if the target axis of the spectrum is folded
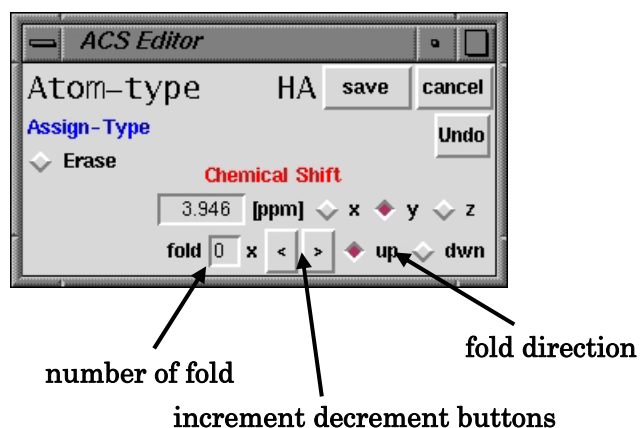
If you find the spectrum on a certain axis is folded by State-TPPI type, you might want to calculate and enter the actual chemical shift of the signal. You can define how many the spectrum is folded and fold direction (up or down field) on the small window, ACS editor as shown the left panel.
Currently, only State-TPPI is supported.

number of fold

increment decrement buttons

fold direction

4) For the assignment of procheral atoms
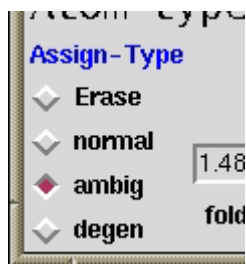
If you would like to enter shift value of procheral atoms such as HB1 and HB2 of Leu, CG1 and CG2 of Val, and you have not achieved stereo specific assignment of the signals, just select "ambig (ambiguous)" check box otherwise select "stereo". If you find the procheral signals are degenerated, select "degen (degenerate)".

5) Amino acids in special chemical states, oxidized Cys, t-His, cis-Pro, and so on.

The ACS module in the new version of Kujira supports cis-Pro, oxidized-Cystein (Cystine) and *t*-His (tautomeric His).   For the Pro residue;
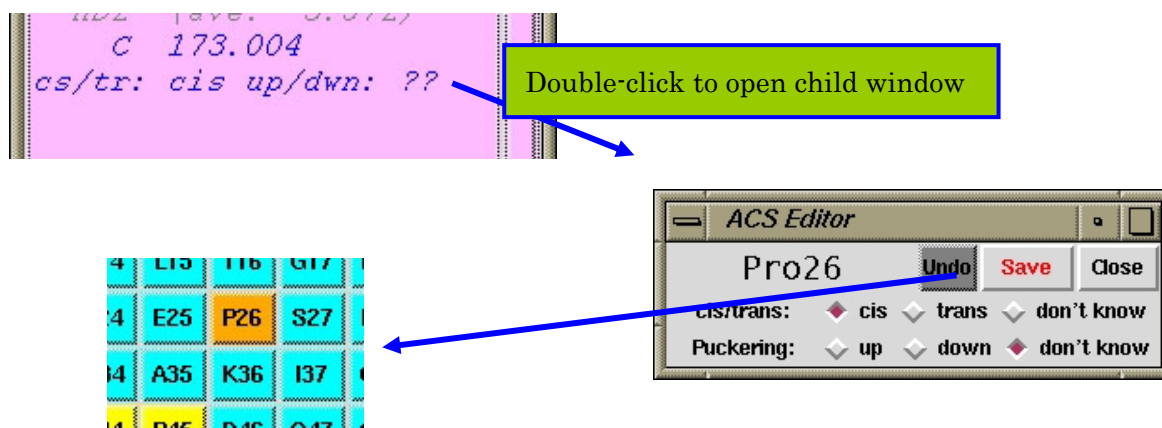
If you defined cis/trans or puckering configuration, the color of the button on the sequence board will change to orange.
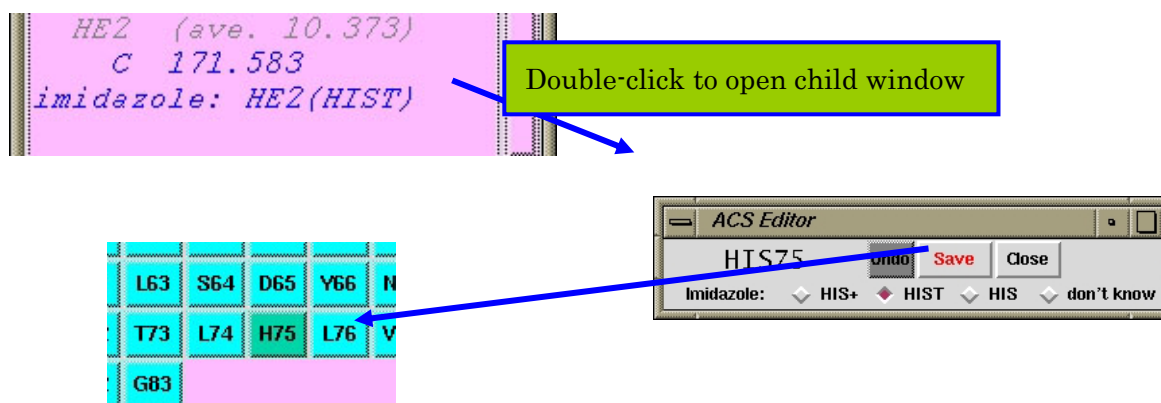
Similarly, for the His residue;

user can select from the four options;

HIS+	fully protonated histidine
HIST	Nϵ2 protonated (so-called tautomeric histidine, t-His)
HIS	Nδ1 protonated

If you defined chemical state of His, the color of the button on the sequence board will change to light-green.

6) If you would like to erase the chemical shift
Check the "Erase" button on the ACS editor.

7) "Acs" directory, in which the chemical shift data you have stored
The shift data you have stored will be found in the directory:

        ~/nv_XXX/matrix/Acs

You can find text files with file extension corresponding to the residue number of the sample.
For example, if you open acs.41 which describes Gly41 chemical shift data, you will see

```
ATOM      560   NH   GLY   41      1.000        8.410
ATOM      561   N    GLY   41      1.000      112.500
ATOM      562   CA   GLY   41      1.000       43.500
ATOM      563   HA1  GLY   41      1.000        4.450
ATOM      564   HA2  GLY   41      1.000        4.450
ATOM      565   O    GLY   41      0.000      999.990
ATOM      566   C    GLY   41      1.000      176.200
```

1: Atom number   Not used by any programs in KUJIRA
2: Atom type   according to the classic format of PDB
3: Residue type   3-letter code
4: Residue number
5: Degeneracy   Used for programs. Degeneracy of the assigned signal
6: Chemical shift(ppm)   Default value is 999.990.

Important: You can edit the text file but please do not forget to take backup of the files before edit them.

Psr (Pattern Search Robot) module

9. Psr, an useful tool for the assignment and confirmation of the assigned signals.

1) What the PSR module is used for?

Psr has a lot of functions to assign and to confirm the chemical shift, signal position, NOEs, ^{1}H-^{1}H distance in the calculated structures and so on. If you press "PSR" button on the green module, you will see a small window as shown in the below panels;
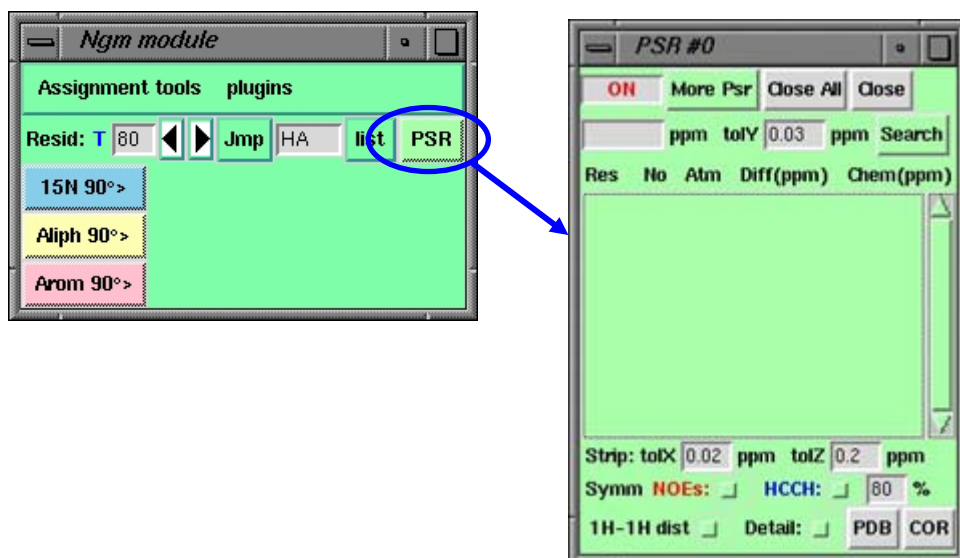
You can open the PSR module as many as possible by pressing "More Psr" button on the module. Note that only one module is sensitive to the chemical shift value insertion. You have to activate each module by clicking "ON" button when you analyze the signals.
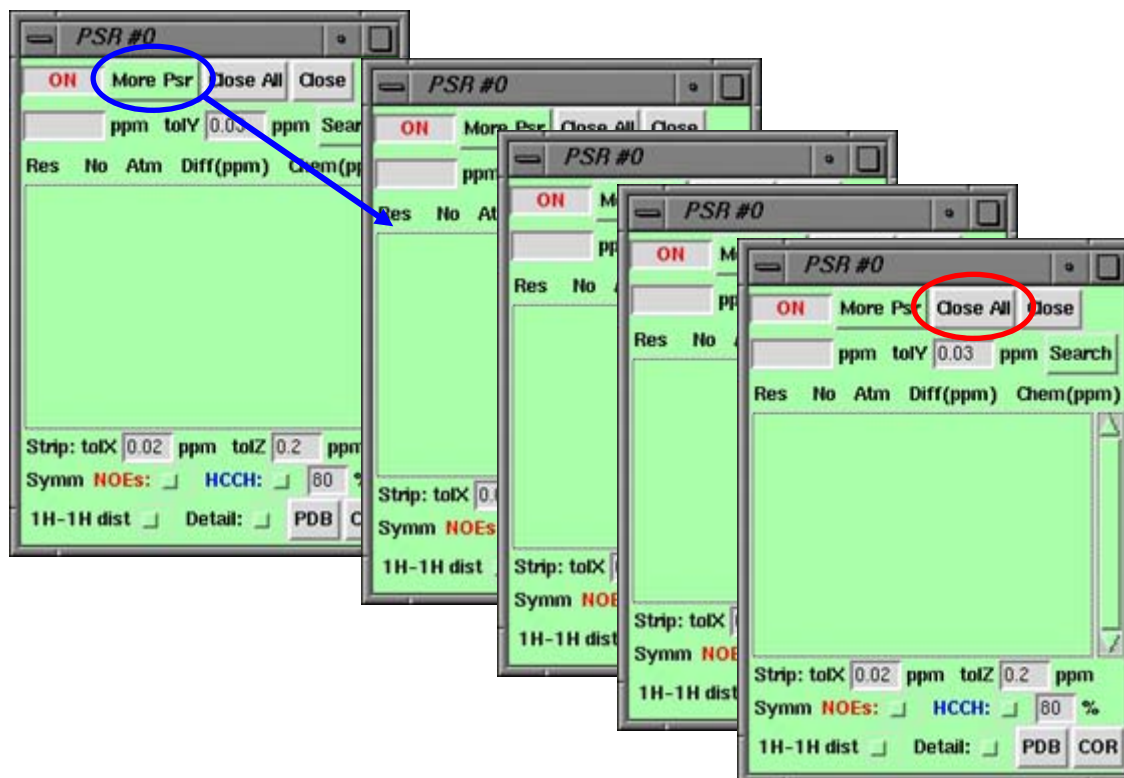
To close all the Psr modules, press "Close All" button on one of the modules.

Psr (Pattern Search Robot) module

2) Signal assignments by the pattern matching search function of Psr module.

If you open a Psr module window, then press "Pattern" checkbox, the window will change to the pattern matching mode.
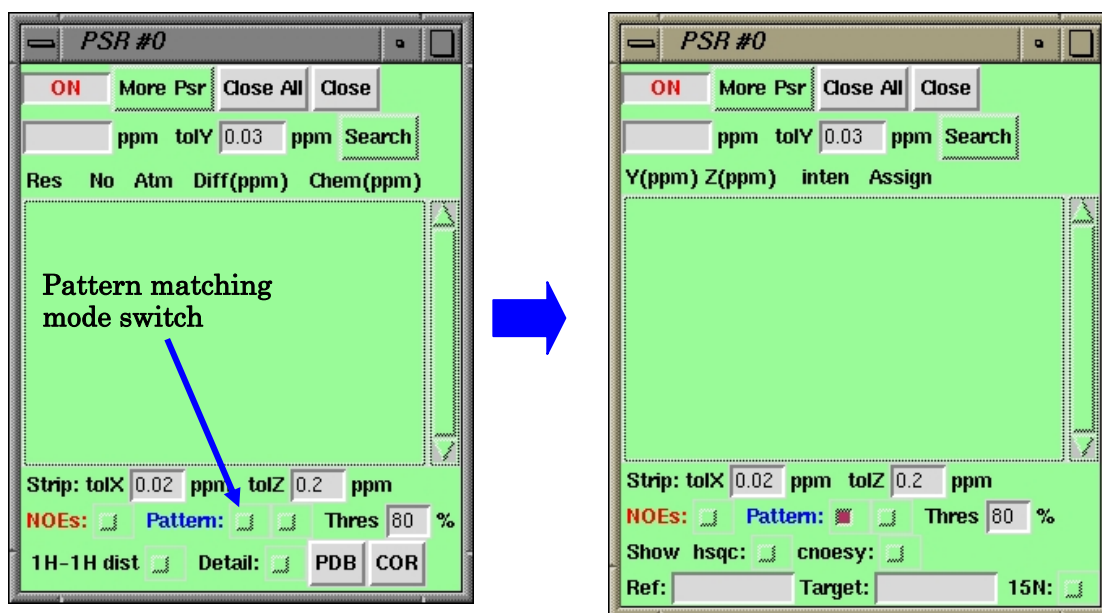
For example you display Hα-Cα spectrum strip of 3D HCCH-TOCSY.   If you would like to address what the signal observed on 1.1ppm is, press middle-mouse-button just on the signal.
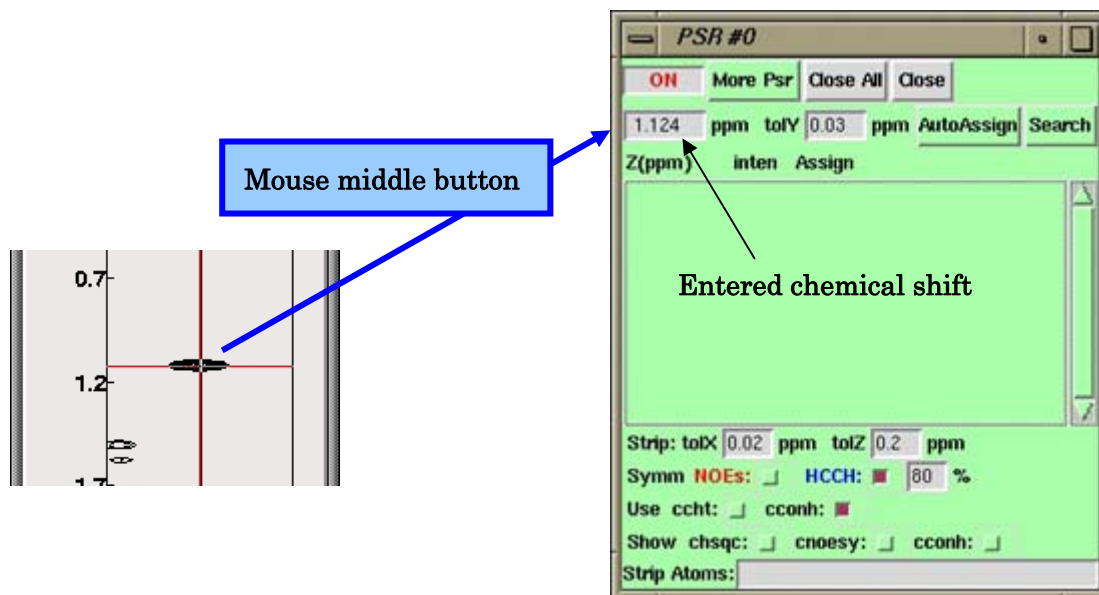
You will see the chemical shift on y-axis has been entered in the module.

Then press "Search" button to search for the possible diagonal signals in the HCCH-TOCSY spectrum.   You can optimize the spectrum threshold and tolerance values for y-, x- and z-axis in the entrance widgets thres, tolY, tolX and tolX, respectively.   The results of the search will be listed on the listbox of the module.

Psr (Pattern Search Robot) module

In the listbox, chemical shift of Z-axis (ppm), intensity and assignment if available are shown for the found peaks. If the program find that the intensity of the found diagonal peak in the corresponding position on the 2D HSQC spectrum, the item will be colored blue.

Hα-Cα   Hγ2-Cγ2
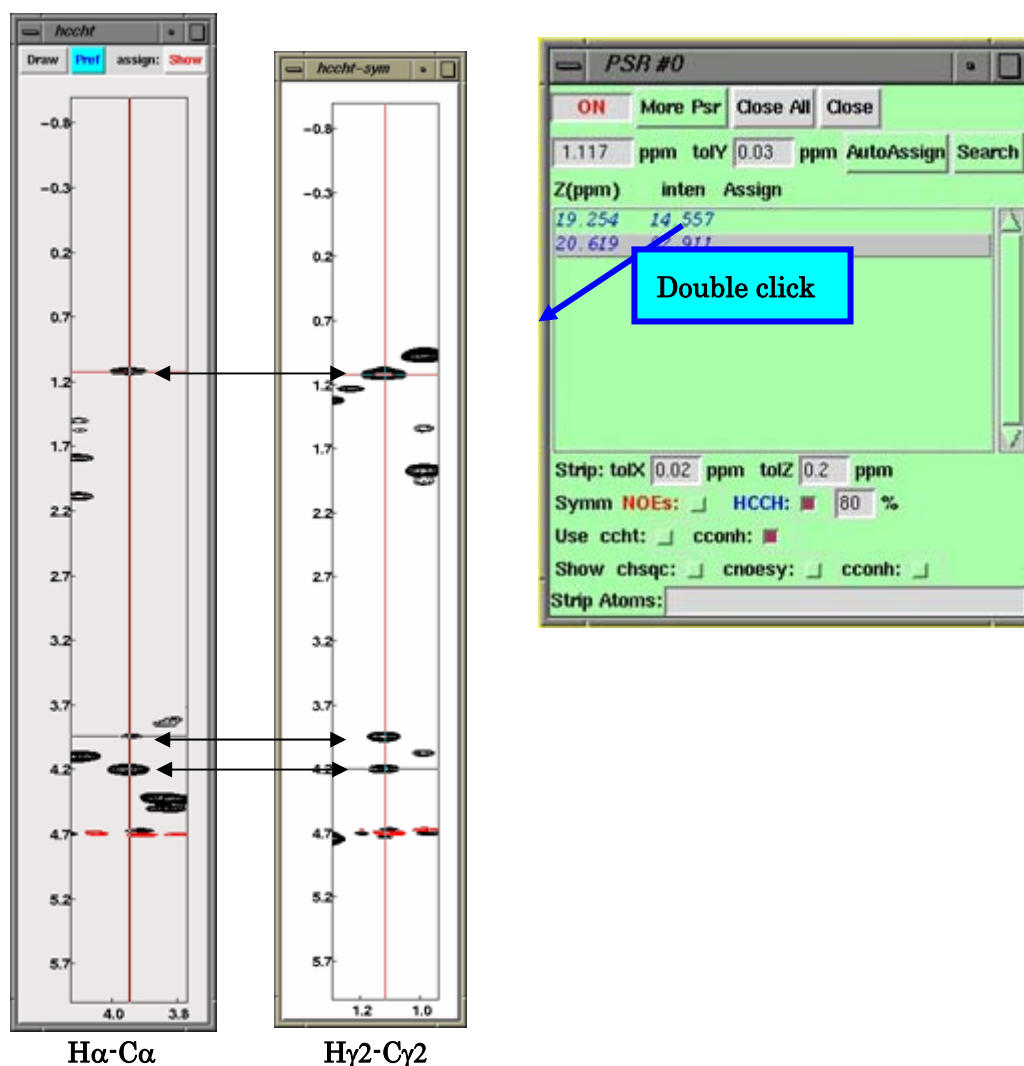
To confirm the found diagonal peak is the real one, double-click on the one of the item shown in the listbox. A whiter spectrum strip will come out to show the spectrum strip of HCCH-TOCSY on the corresponding spectrum position of the found peak.
If you find the peak patterns are very similar and the carbon chemical shift is reasonable, store the assignment in the magenta Acs module.

Psr (Pattern Search Robot) module

3) The NOE assignment of y-axis direction for 3D NOESY

If you place the cross-hair on the NOE signal as shown below, then press mouse-middle button, the chemical shift on the cross-hair position will be entered in the module.
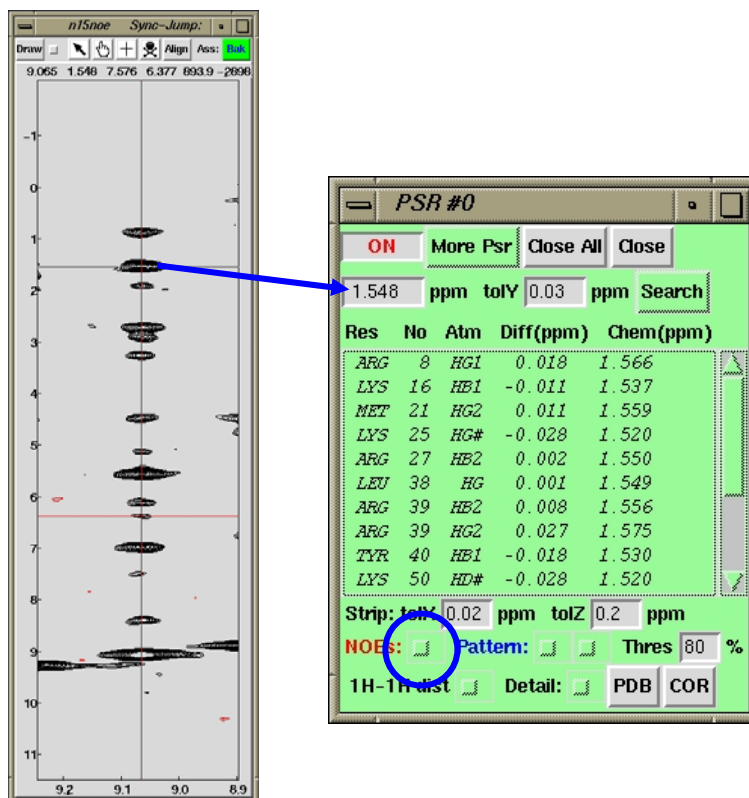
Then press "Search" button to search for the corresponding proton signals in Acs chemical shift table within the error specified in another entry "TolY". The found protons will appear in the listbox sorted by the chemical shifts.

If you activate the checkbox on the module as indicated in the above panels with the blue circle, the module
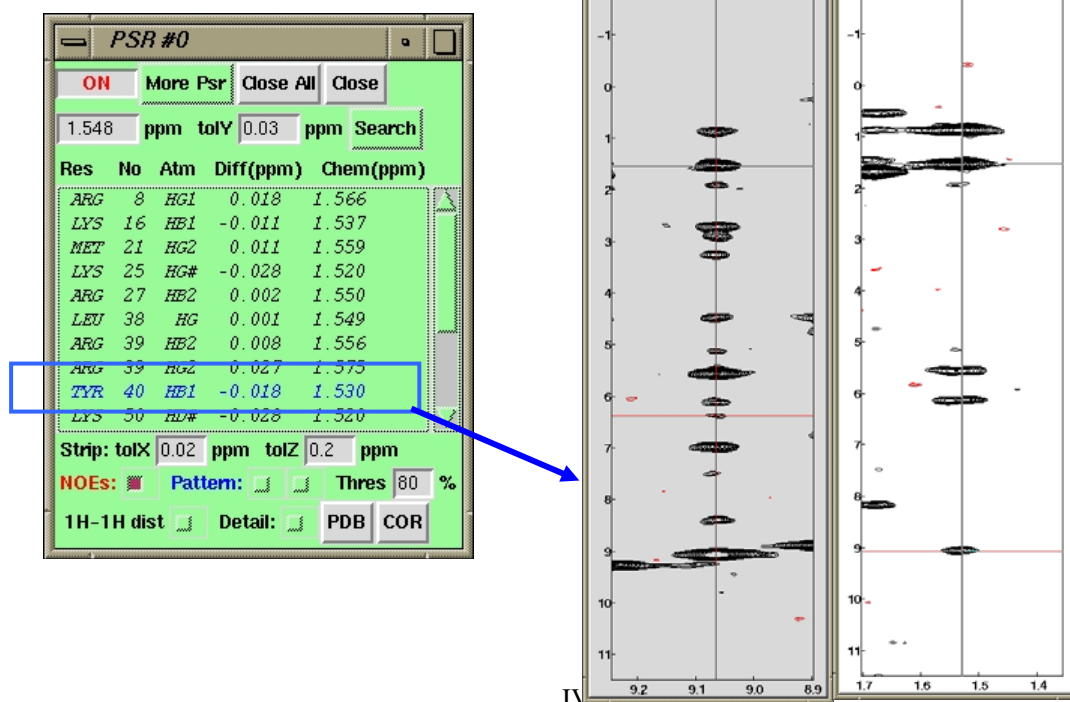
Psr (Pattern Search Robot) module

will also examine the symmetry of the NOE assignments by checking the spectrum intensity on the corresponding transposed position.   If the NOE assignment symmetry has been confirmed, the item shown in the list will be colored by blue.   If you would like to see the spectrum strip of the transposed position, double-click the item to generate new 2D-spectrum strip showing the spectrum region on the corresponding transposed position.

4) ^{1}H-^{1}H distance calculation mode of PSR module

If you press the check button on the module "1H-1H dist", the small module will change to the mode "^{1}H-^{1}H distance calculation mode" as shown in the left panel.   In this mode, the module can calculate and display the actual ^{1}H-^{1}H distance in the loaded protein structure for the assigned NOEs.
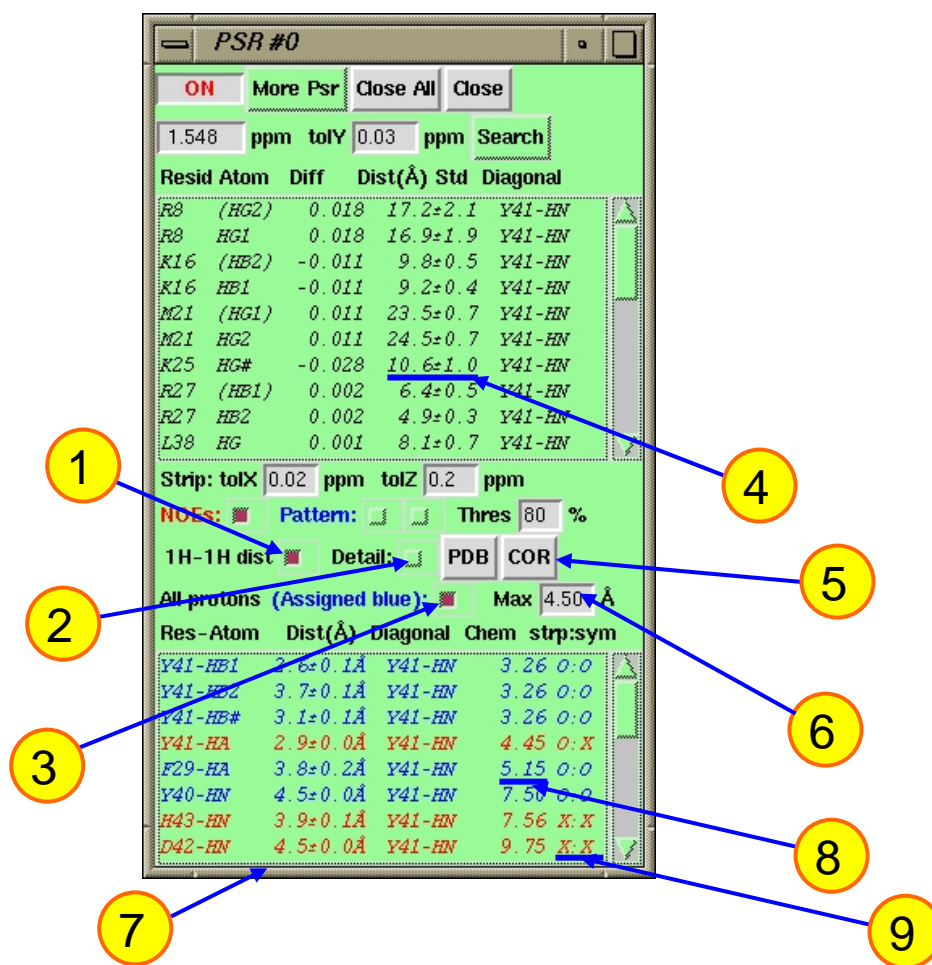
1: Switch the mode of PSR odule to ^{1}H-^{1}H distance calculation mode

2: Switch the module to the mode for showing all the detail of the distance calculations.

3: Switch the module to the mode for displaying all pair of protons: between all of the protons expected to be observed in the current spectrum strip and the protons involved in the specified radius in the entry 6.

4: Calculated ^{1}H-^{1}H distances (Å)

5: Load protein structure file in PDB or COR (CYANA) format

6: Maximal distance to display the all pair of ^{1}H-^{1}H distances for the function 3

7: The list box showing the calculation results performed by the function 3.

8: The chemical shifts if the found protons have been assigned in the Acs.

9: The intensity check result on the protons on the current spectrum strip and symmetrical (transposed) position of the NOE assignments.

Psr (Pattern Search Robot) module
5) Y-axis assignment for the other type of atoms on PSR module
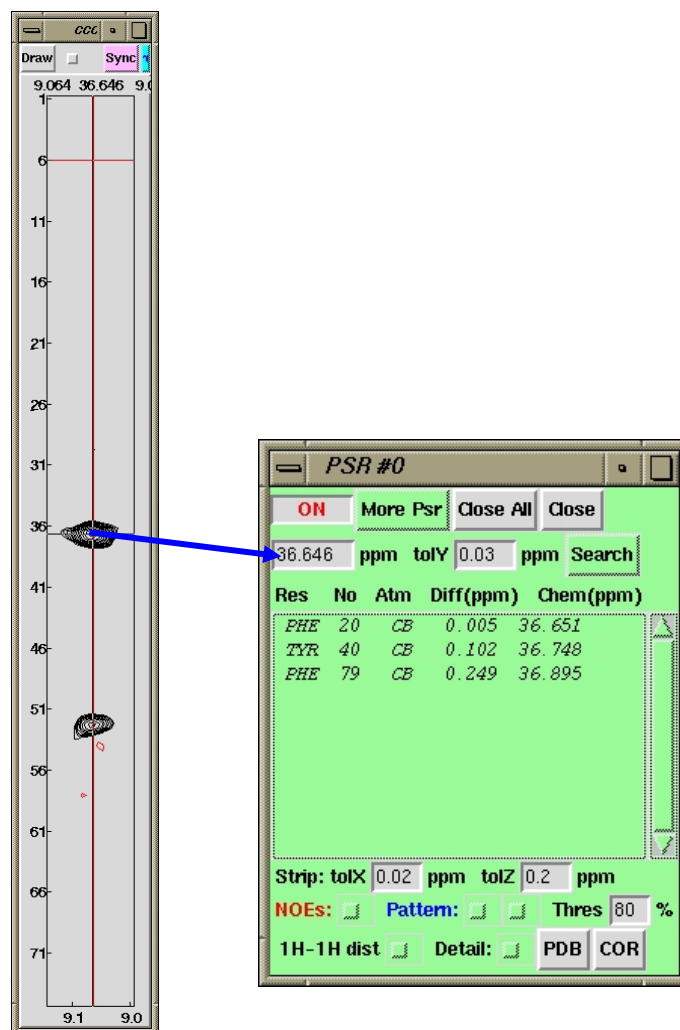
PSR module can be used for the assignments of ^{13}C and ^{15}N atoms.

The left panel demonstrates the y-axis signal assignment of the 2D spectrum strip of C(CCO)NH on y-axis. The ^{13}C signal assignment has been performed for the chemical shift detected from the 2D spectrum strip by mouse-middle-click based on the assignment table "ACS" with the search tolerance 0.3ppm. Note that, although the value "tolY" is set at 0.03ppm, the used search tolerance was 0.3ppm (automatically multiplied by 10).

10. Confirmation of side-chain signal assignment using "Show strip" function on the magenta module

Requirements: 3D ^{15}N-edited NOESY, ^{13}C-edited NOESY, HN and ^{15}N assignments

If user have loaded 3D ^{15}N-edited and ^{13}C-edited NOESY spectra, and at least HN and ^{15}N signals have been assigned, user will be able to use "Show strip" function. If you press "Show Strip" button on the magenta module, a large window will open as shown in the below panel.

^{15}N-NOESY

^{13}C-NOESY aliphatic

Functions of the window

1: Button to start "Fully automated one residue assignment" (requires AssignRobot)

2: Residue number entry, **3:** Increment and decrement button for residue number

4: Entry for corrected shift value **5:** Save button for storing corrected shift value

6: "Show" button for displaying assigned signal position on the spectrum strips with blue boxes

7: "none" button for removing assign boxes displayed on the strips.

8: 90 degree flip on y-axis button

Confirmation of side-chain signal assignment

The "Show strip" function will display two dimensional ^{1}H-^{15}N or ^{1}H-^{13}C spectrum strips corresponding to the assigned signal of the current residue number. ^{1}H-^{15}N and ^{1}H-^{13}C HSQC NOESY spectra are used for ^{1}H-^{15}N and ^{1}H-^{13}C strips, respectively.

If you find that a certain chemical shift is wrong, you can correct the value manually on the window.

For example, if shift value of HG1 methyl signal seems to be wrong, place cross-hair on the signals that are expected to be NOEs from HG1 methyl. Then press mouse-middle button, the x-axis value will enter the entrance on the window header. Atom type will be automatically recognized. To store the corrected value, press "Save" button. If you would like to confirm the corrected value, press "Show" button to redraw the spectrum strips and assignment boxes.

mouse-middle button

Press "Save" button
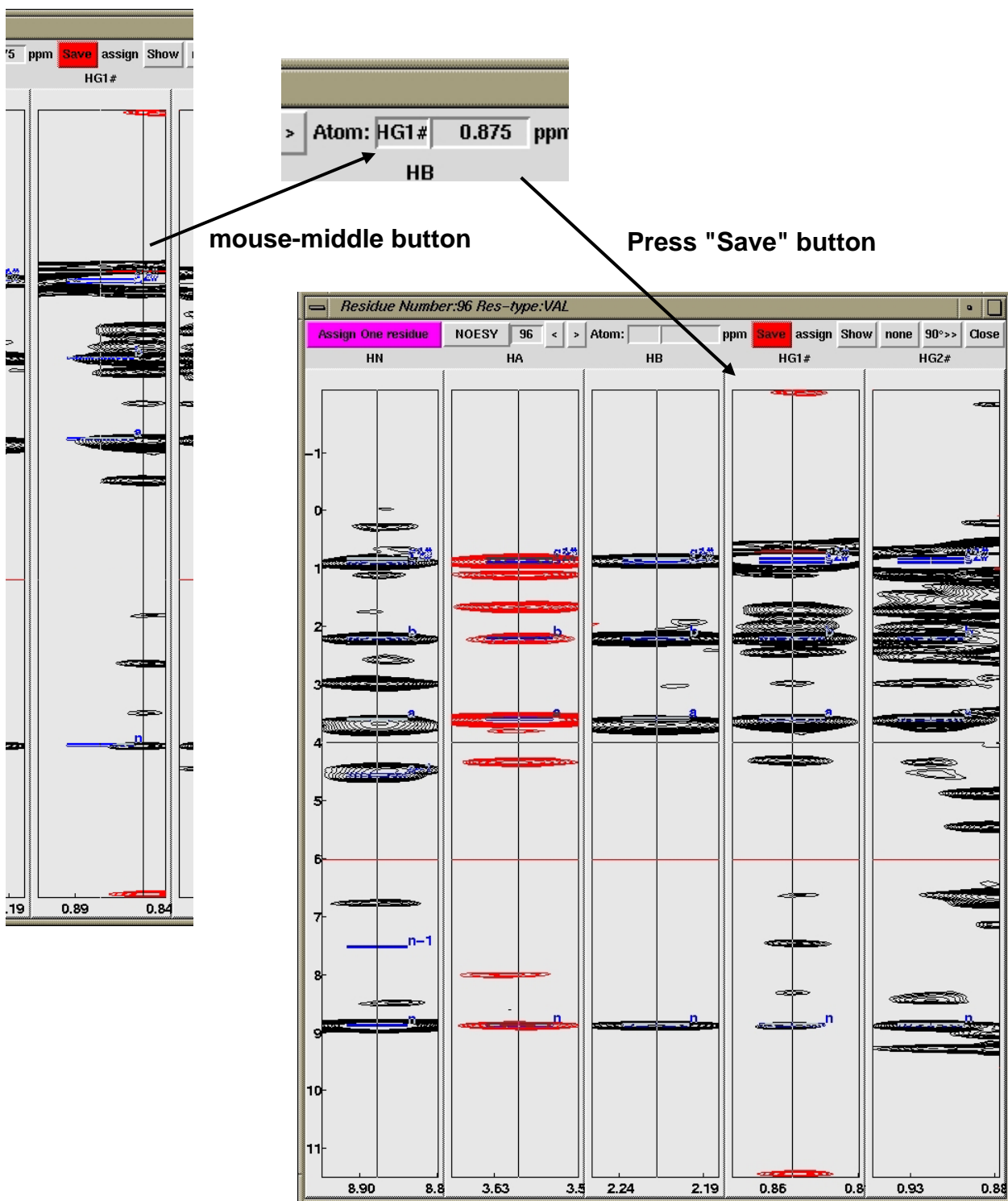
CYANA setting module

11. CYANA setting module, a semi-automated input file maker for CYANA calculations

The module for creating CYANA calculation input files is now available from the magenta module (ACS). The settings are mostly automated, user will be released from the annoying works such as conversion of file format, checking version dependency, and so on.

1) Requirements

At least one NOE peak table, 3D-NOESY and ACS directory

<out-put files>

|	
CALC.cya, int.cya	defines CYANA calculation parameters such as constrain file names, violation tolerances, library names, stereo chemistry settings, etc.
*.seq	sequence file. Can be specified cis-Pro, Cystine and t-His **(library should support them).**
*.prot	chemical shift table (now support only one *.prot file).
talos.aco	phi, psi angle constraints derived from TALOS prediction **(pred.tab required)**
*.peaks	supports ^{1}H-^{1}H-^{15}N and ^{1}H-^{1}H-^{13}C xpk files.

[Important]

Currently, this module supports CYANA ver. 1.0.x beta, 2.0.x beta and official release ver. 2.x.   At least one *.prot file and two *.xpk files derived from 3D ^{15}N-edited NOESY (axis order ^{1}H-^{1}H-^{15}N) and 3D ^{13}C-edited NOESY (axis order ^{1}H-^{1}H-^{13}C) are required.

CYANA setting module

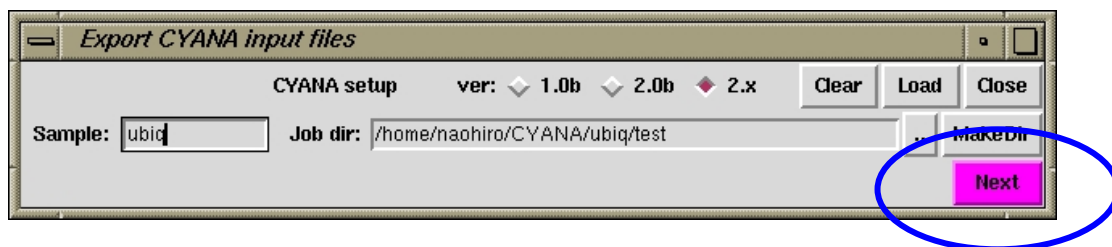## 2) Quick start CYANA input file maker

If you click the button "Plugins" on the left-top of the magenta module, you will see a pull-down menus;,
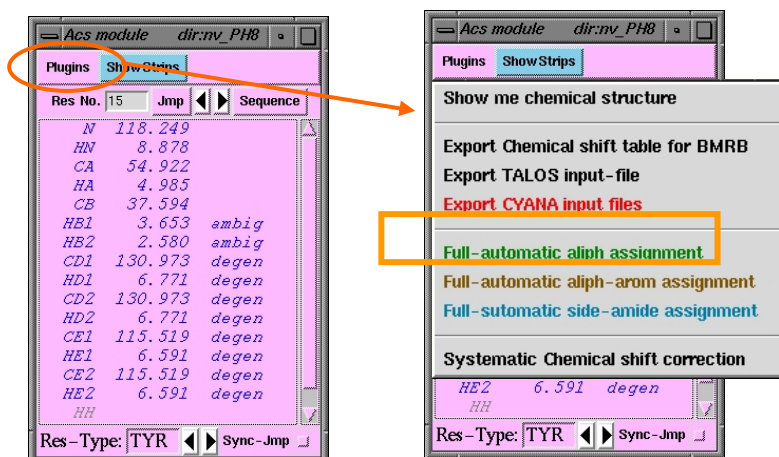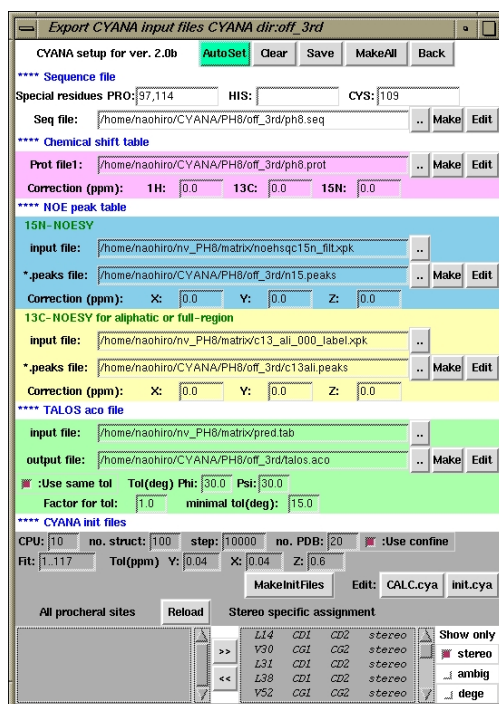

Select "Export CYANA input files" among the menus to open startup window;

Type the sample name and the path name for CYANA calculation and select version of CYANA as shown above panel.

Then press "Next" button to open the main-setting window

CYANA setting module

The initial state of the setting entries are empty.   **If you don't know what to do**, please try to press "AutoSet" button, the automated setting program will fill the entries.
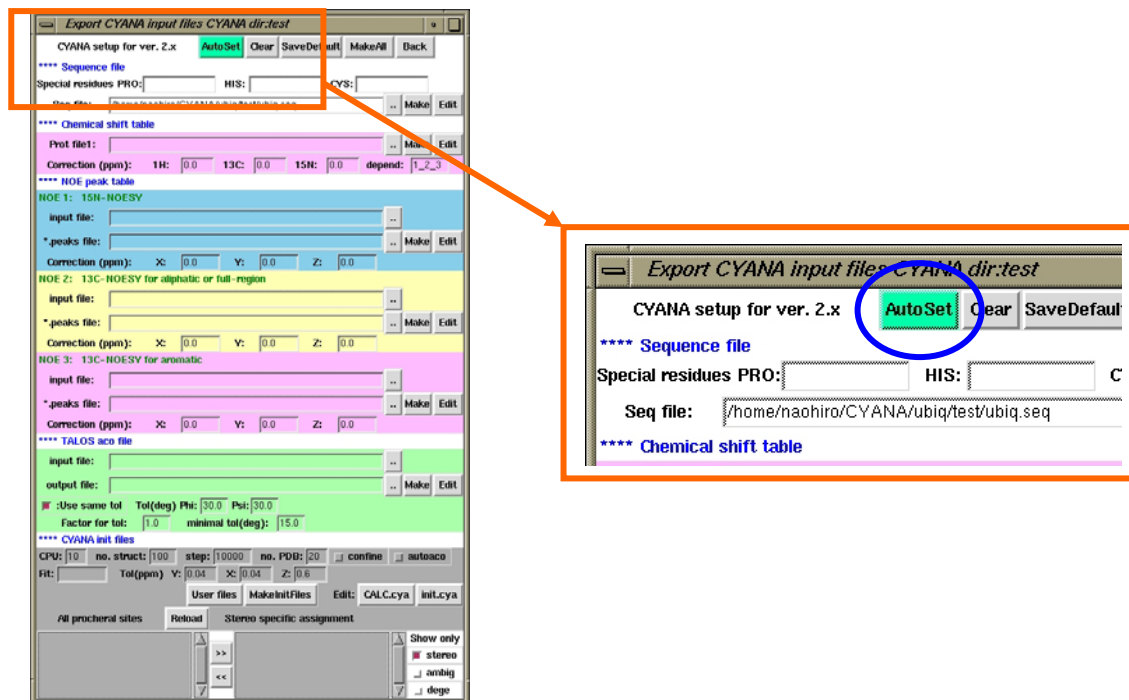
If the settings seem to be OK, then press "MakeAll" button (blue circle) to make all input_files.

You can also make each input_file by pressing "Make" button (green circle) on each file setting section.

And if you would like to edit the created file, click "Edit" button to open the text file editor

By pressing the "SaveDefault" button, you can save the current setting as default for the next calculations ,

CYANA setting module

3) Details of the input file sections

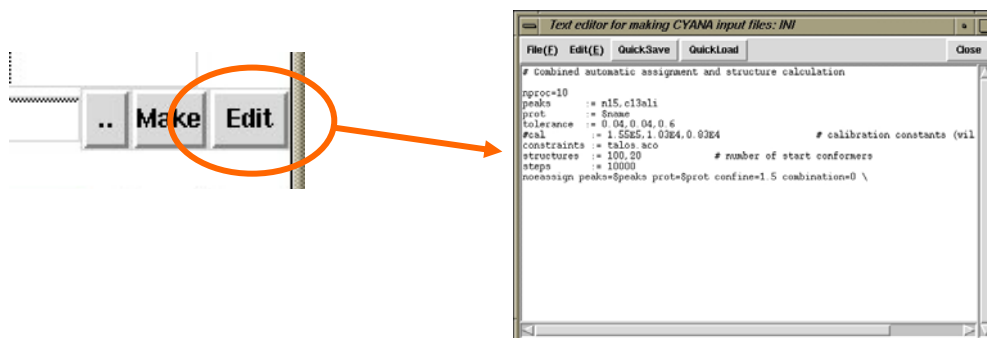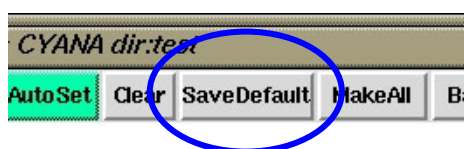**A) Chemical shift table:**

For the chemical shift table file, *.prot, the GUI stuffs are;

The chemical shift values are loaded from the ACS files, matrix/ACS/acs.*.   The ^{15}N and ^{13}C chemical shift values for the signals on the folded region of 3D-NOESY spectrum are calculated using the NOESY spectrum parameters if the spectrum dependencies are defined in the entry "depend:".   The example shown above indicates, ubiq.prot file <--> NOE:1 and NOE:2.

B) NOE peak table:

For the NOE peak table file, *.prot, the GUI stuffs are;

The most recently edited xpk file is automatically found and typed in the "input file" entry.   User has to define the name and path of the output-file in the "*.peaks file" entry   Currently user can specify three NOE peaks files, ^{1}H-^{1}H-^{15}N NOESY, ^{1}H-^{1}H-^{13}C NOESY for aliphatic and aromatic regions.

CYANA setting module

C) CALC.cya and init.cya files:

The GUI stuffs for making CALC.cya and init.cya files,are found on the bttom gray zone of the module as shown below.
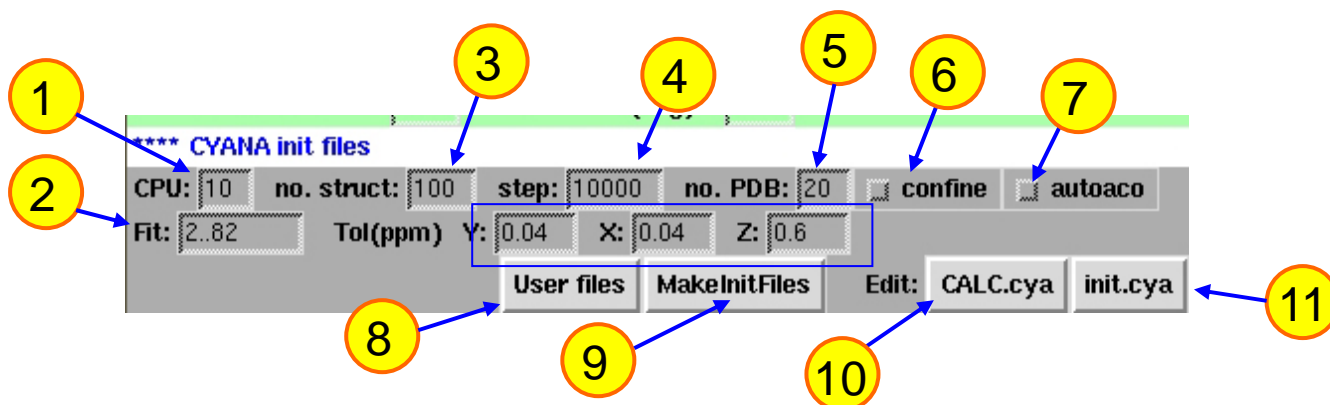
For the settings for calculation;

      1:      number of CPU for the CYANA calculations. [default 10]

      2:      residue numbers for fitting calculated structures. [default begin-residue-1 - end-residue]

      3:      number of structures for calculation [default 100]

      4:      number of steps for calculation [default 10000]

      5:      number of selected structures [default 20]

      6:      confine mode distance constrain [default disable]

      7:      automated aco file making [default disable]

      8:      call child module for setting user defined files for CYANA calculations

      9:      button to make CALC.cya and init.cya files

    10 & 11: buttons to call for text-editor windows
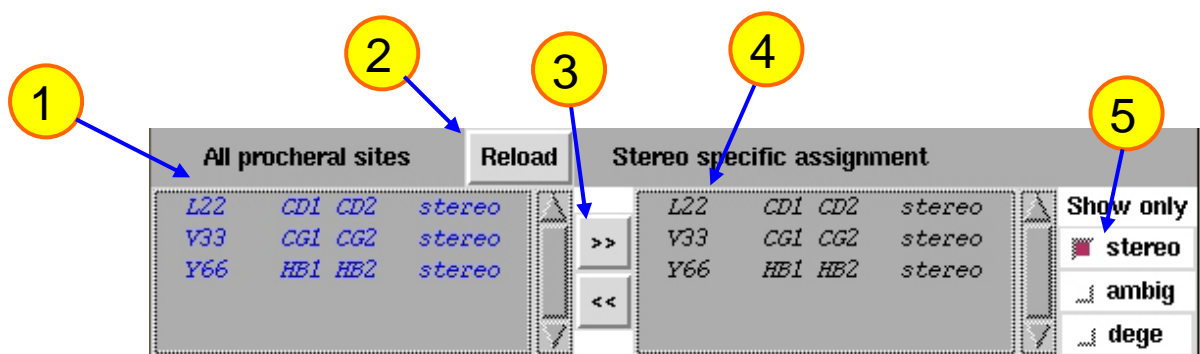
And for the stereo chemistry settings;

      1:      left listbox to show the procheral atoms found in the chemical shift table, ACS. The atoms colored by blue indicate the stereo-checmistry is defined to include CYANA calculations

      2:      reload the stereo-chemical atoms from ACS

      3:      add and delete the elements listed in the right-side listbox

      4:      right listbox to define the stereo-specific assigned atoms used for CYANA calculations

      5:      If you select "stereo", only stereo-specifically assigned atoms are loaded when you reload ACS.   If you select "ambig", the atoms whose stereo-chemistry is ambiguous are loaded. If you select "dege", all procheral atoms are loaded.

CYANA setting module

4) If you want to use custom files that you have made such as *.upl, *.lol, *.aco and *.lib files.....

     User is allowed to specify the files for additional constrain files and library files.   If you click the button on the section for setting CYANA input files (red circle), you will see a child window.   You can define the file name and path in the listbox for the following files;

        *.upl   file for upper limit distance constraints
        *.lol   file for lower limit distance constraints
        *.aco   file for dihedral angle constraints
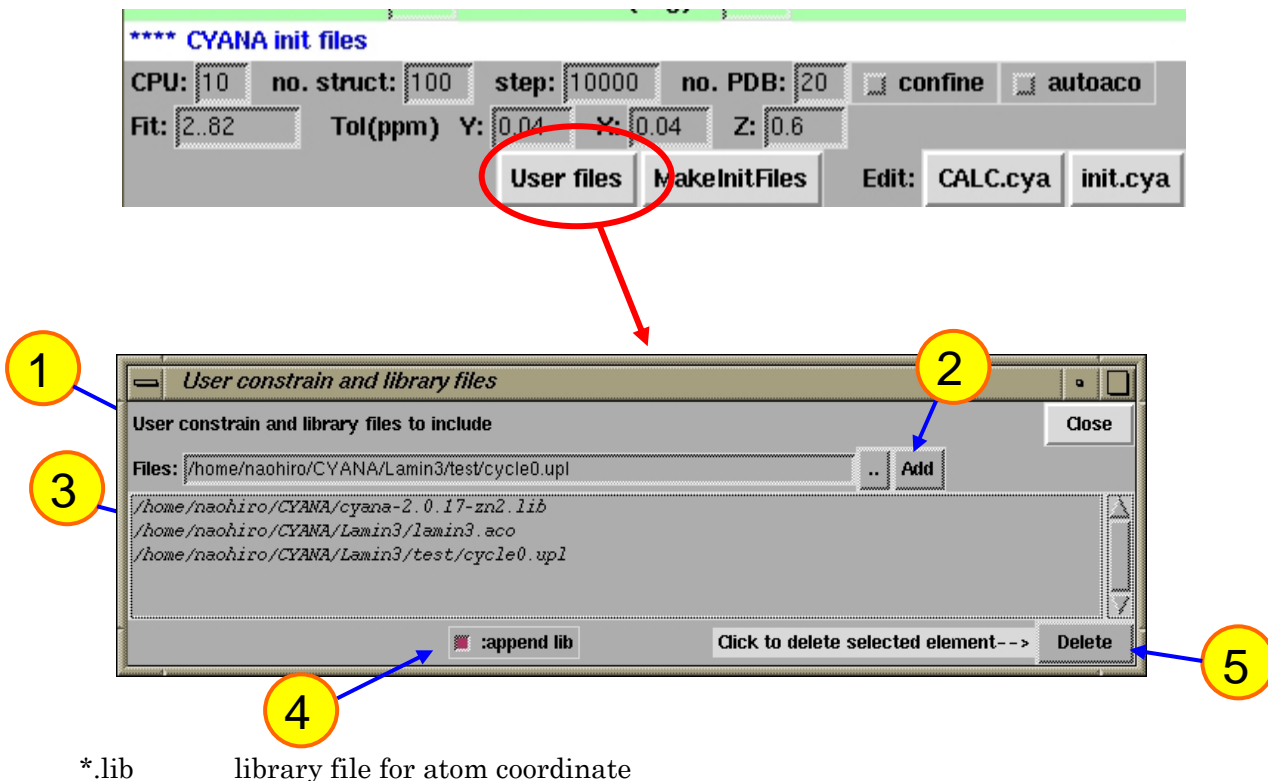
      *.lib   library file for atom coordinate

.

          1:   entry to define the path name of the target file
          2:   add button used to add the defined path name to below listbox
          3:   listbox showing files going to be used for CYANA calculation
          4:   append-lib checkbox.   If it is checked, the listed library file will be appended to the standard library file for CYANA calculation.
          5:   delete button to delete selected item in the listbox

If you click to execute "MakeInitFiles" button or "MakeAll" button, the CALC.cya and int.cya files are created with these informations.

12. CYANA result analysis module

1) Setting up the module
The "CYANA result analysis module" finally supports the official version of CYANA, 2.x.

<Requirements>
        CYANA input files and output files;
        CALC.cya and **init.cya** files
        cycle*.noa and **cycle*.upl** files
        Spectrum data files, *.nv, for 3D ^{1}H-^{1}H-^{15}N NOESY, 3D ^{1}H-^{1}H-^{13}C NOESY

An example of the CYANA result analysis shown in this page has been performed with the following data sets;

| Sample:	ubiquitin	
	NOESY:	3D ^{1}H-^{1}H-^{15}N NOESY, 3D ^{1}H-^{1}H-^{13}C NOESY for aliphatic and aromatic
	NOE peak table:	n15.peaks and c13noe.peaks obtained from the NOESY spectra
	Chemical shift table:	ubiq.prot, refined with the NOESY spectra
	CYANA:	official release ver 2.1
	target NOA:	final.noa

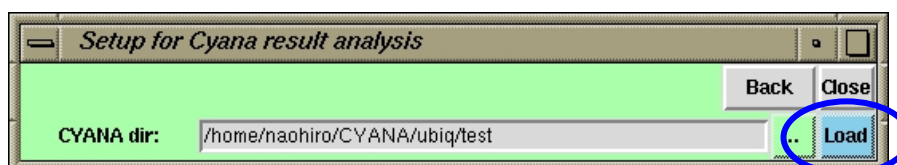To start the setting up the CYANA result analysis, click the "plugins" button on the green module (SSM) to open
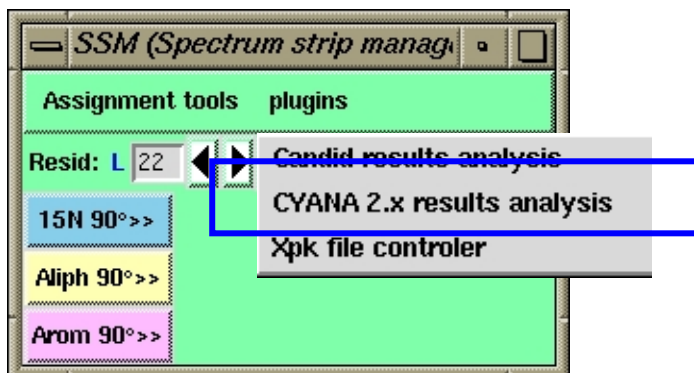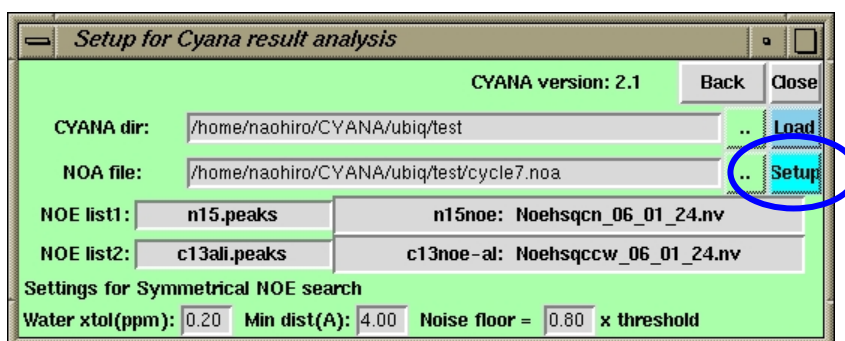

the pull-down menus.
Select "CYANA 2.x results analysis".   Then you will see an initial setup window
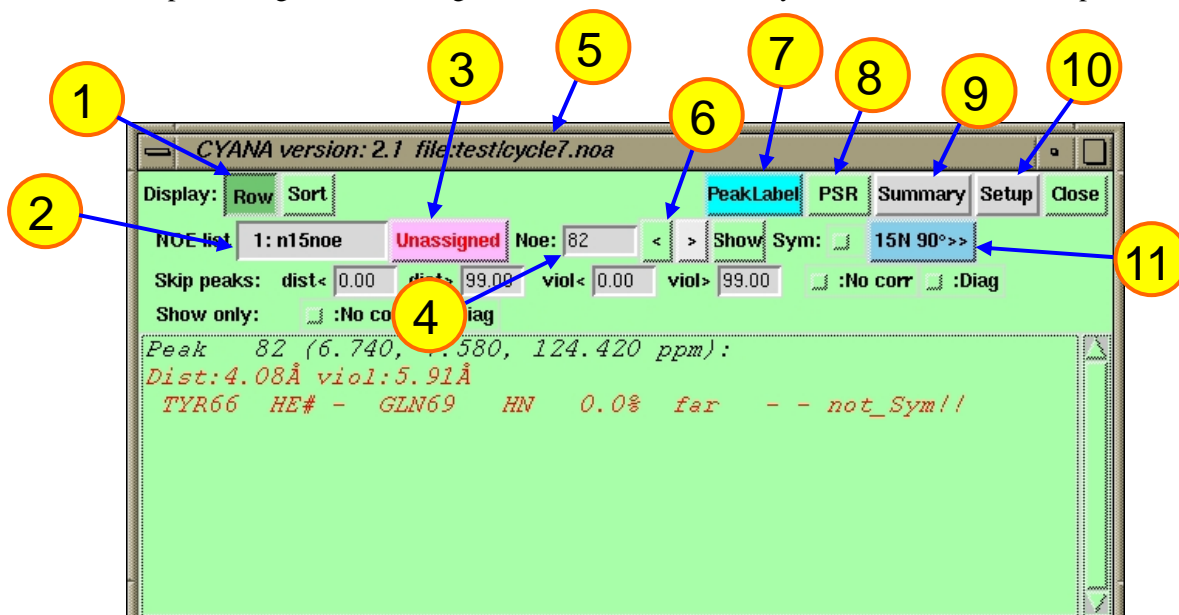Type the path name of the CYANA calculations, then press "Load" button to load the information.

The window will change to show additional GUI items;

Press "Setup" to finish the setting up jobs.

CYANA result analysis module

2) Functions of the main window

1) Row mode.　　Representing the NOE assignment results carried out by CYANA for each NOE peaks.

1:	Button to switch the window between "Row" and "Sort" modes.
2:	Pulldown button to switch NOE peak lists
3:	Button to switch between "Assigned" and "Unassigned" NOE peak analysis modes
4:	Entry to set NOE peak ID
5:	Window header displays version of CYANA and the current path name of *.noa file
6:	Increment and decrement buttons for NOE PeakID number
7:	Button to open setting window for peak labeling
8:	Button to open PSR module
9:	Button to open Summary window
10:	Button to open main-setting up window.
11:	Button to execute 90deg flip of the NOESY spectrum strip
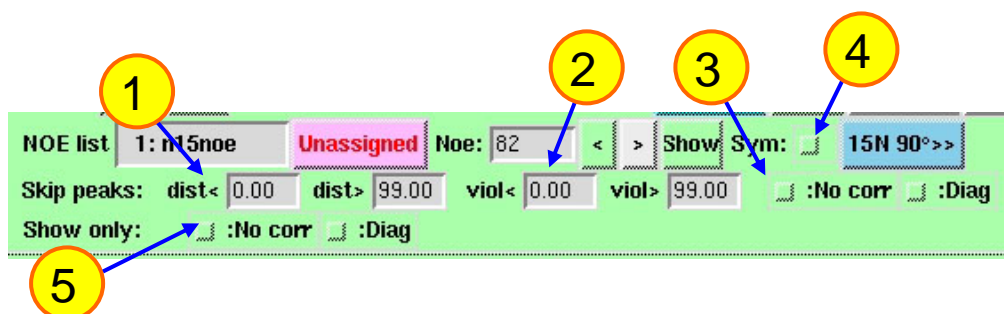
And options to select PeakIDs are;

1:	options to skip NOE peaks within the defined distance range.
2:	options to skip NOE peaks within the defined violation range
3:	option to skip peaks of "No corresponding signal" and diagonal peaks
4:	switch to enable the **"twin strips display"** to check NOE assignment symmetry
5:	option to exclusively show "No corresponding signal" and diagonal peaks

CYANA result analysis module

3) The Sync-Jump from CYANA analysis module

The module can provide "Sync-Jump" command to control NOESY spectrum strip corresponding to the selected NOE peakID.   By pressing "Show" button, the target 3D-NOESY spectrum get jumped to the peak positions   "<" or ">" button increments and decrements peakID number as well as jumps the target NOESY spectrum.   If checked the option "Sym" (red circle), the module turns to "twin strips display" mode.   The "Sync-Jump" command displays the spectrum strip corresponding to the transposed peak position of the NOE assignment.   The function is useful to confirm the **symmetry** of the NOE assignment achieved by CYANA.
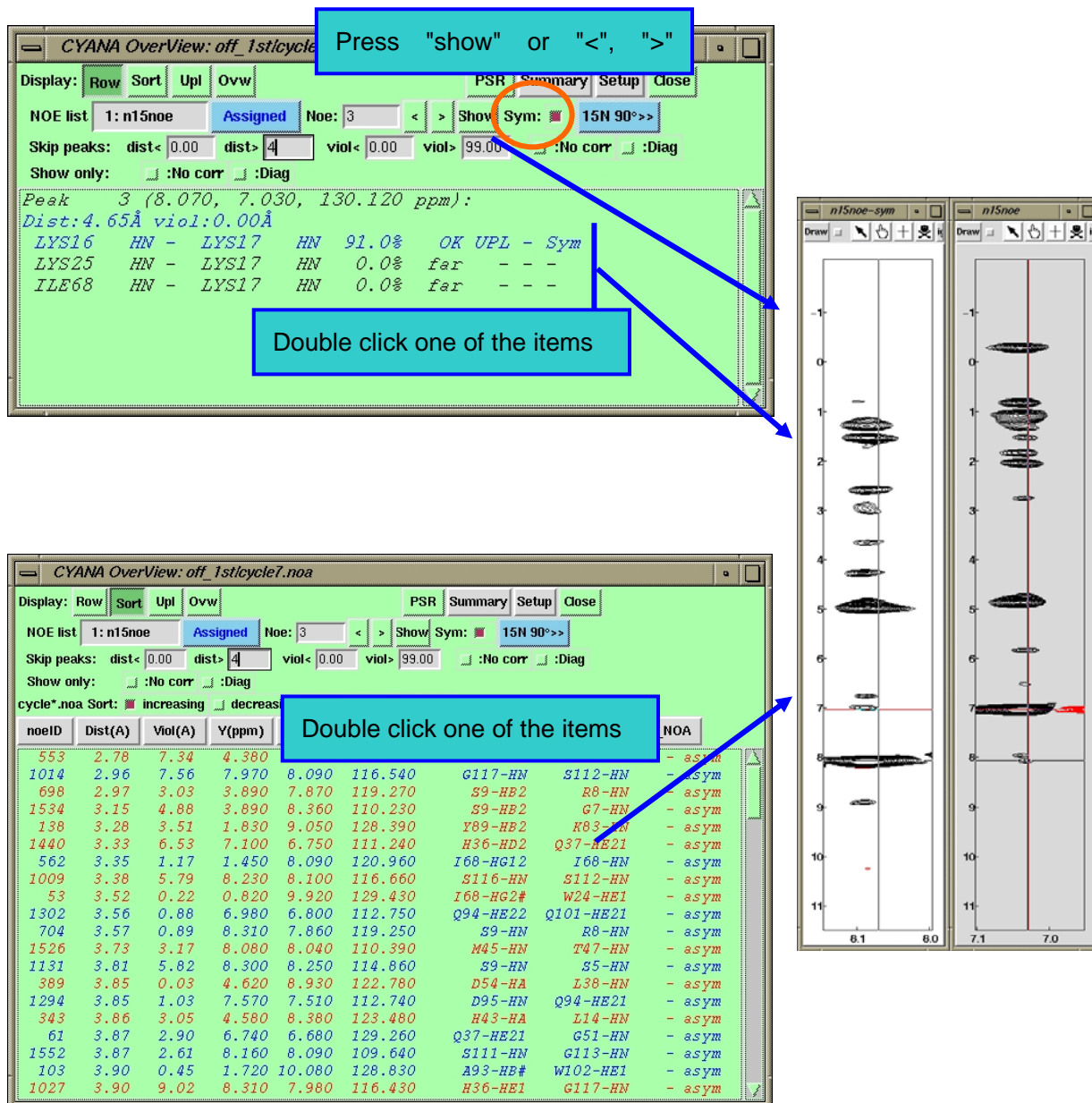
CYANA result analysis module
4) Sync-Jump of spectrum strip with fixed region.
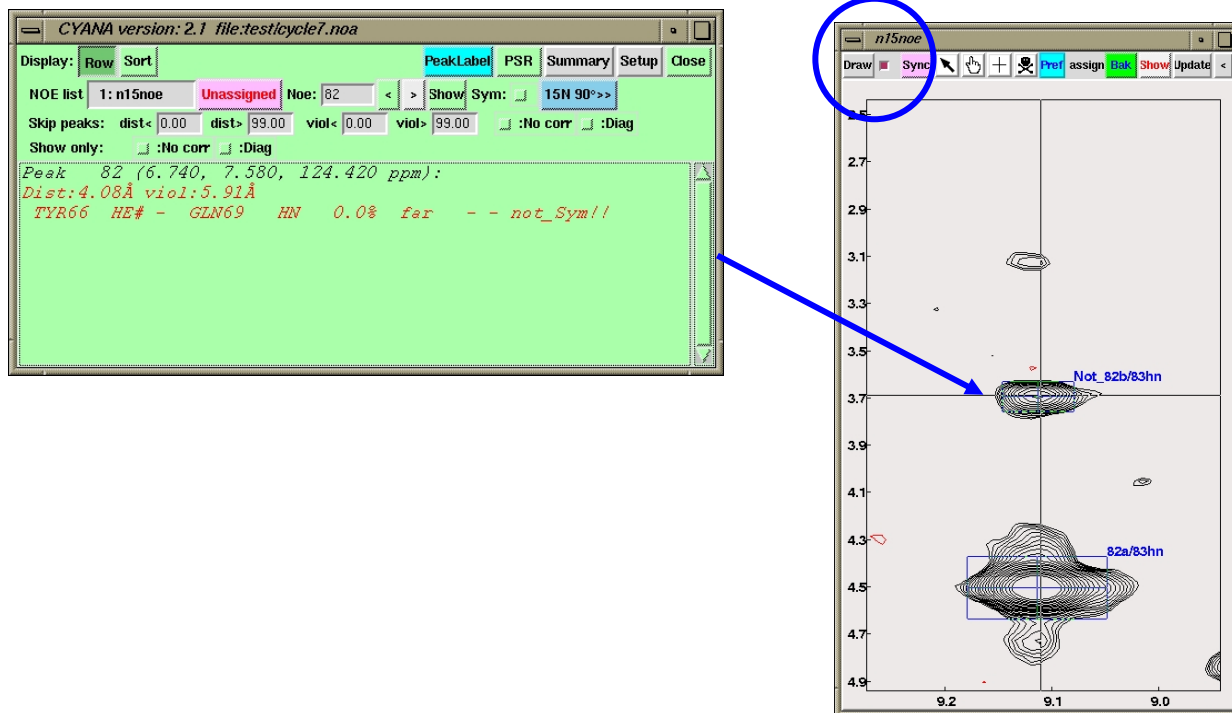
A small checkbox can be found on the every 3D-spectrum strip. By activation of the checkbox, the "Sync-Jump" command from CYANA result analysis module controls spectrum strips with fixed window apspect. This function is usefult to recognize NOE peaks at desired spectrum aspect ratio.

CYANA result analysis module

5) Peak labeling function

If user press "Peak label" button, the setting window will appear.

To run the program, user has to define input and output xpk file names, and then press "Execute" button to finish the labeling job.
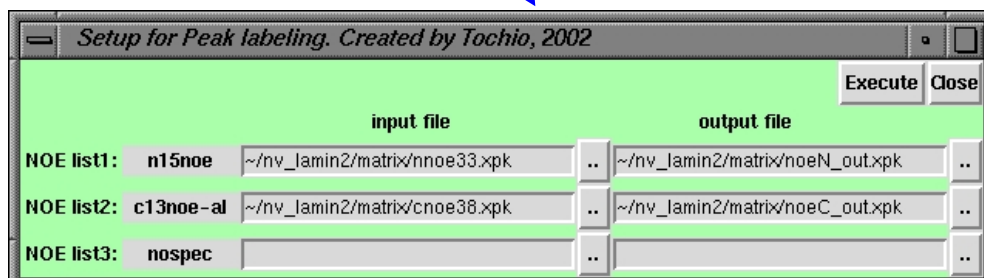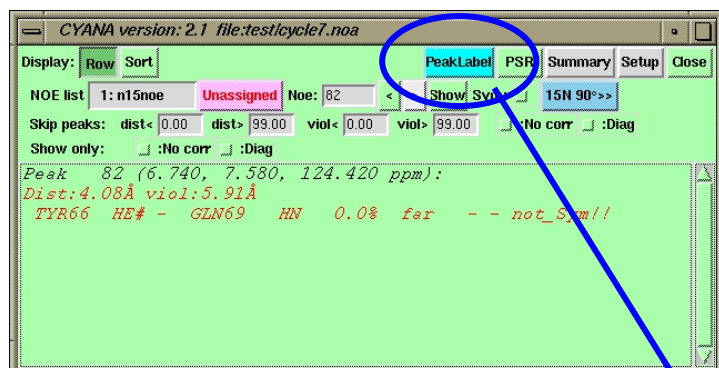
The peak labeling job finishes in 1-2sec for each xpk-file (including ~200 peaks).

The following are what the labeling codes indicate;

Diagonal peak	Xdiag
No corresponding signal	X0/0
Not assigned, Ala32-Ha <-> Thr55-Hg#	X32a/55g#
Assigned, Met103-Hg# <-> Gln29HN	103g#/29n
Assigned intra-residual, Phe67-HB2<->Phe67-HD#	b1/67d#

13. Structure assessment function of KUJIRA

1) How to use structure assessment tools on ACS module

The structure validation tools are implemented in the ACS module.   If user clicks the checkbox "structure" (blue circle) on the header of ACS module, a child listbox appears on the bottom of the ACS module window.   Before start to use the validation tools, user has to define structure file by clicking the "PDB" or "COR" buttons (red box) respectively for *.pdb or "*.cor" file.
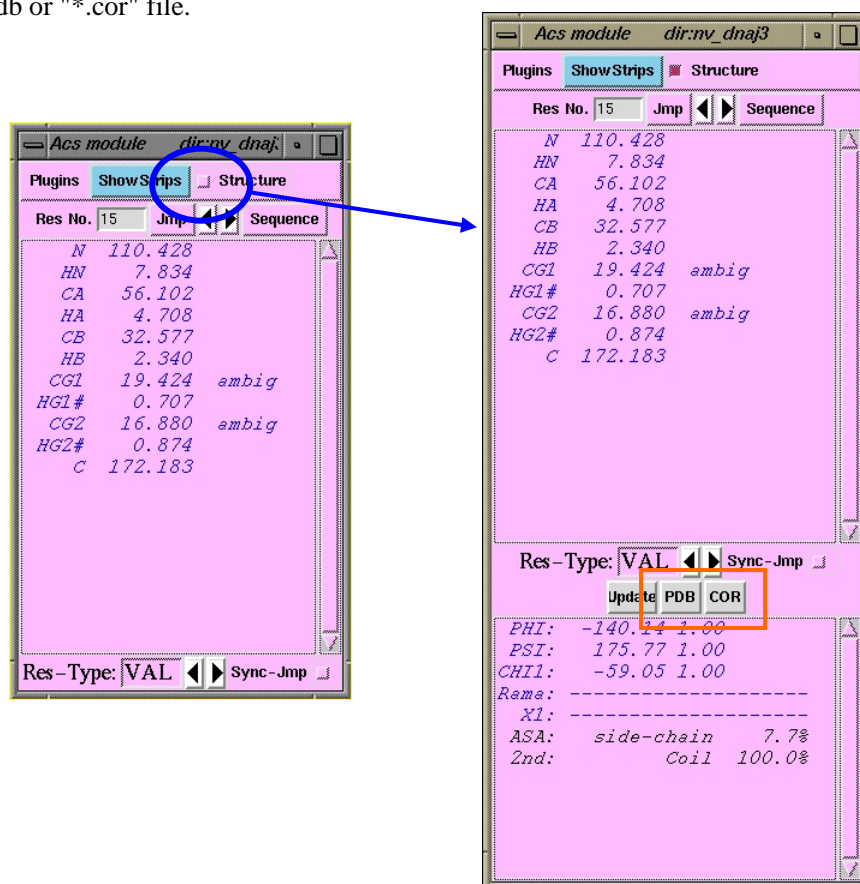
The child list box displays

PHI, PSI:	the averaged values and order parameters for $\phi, \psi, \chi1$ and $\chi2$ dihedral angles,
Rama:	ϕ and ψ values on the 2D-Ramachandran plot analysis
X1X2:	$\chi1$ and $\chi2$ values compared with those found in statistic rotamer library
ASA:	the averaged accessible surface area of the side-chain atoms
2nd:	detected Secondary structure

For the analysis of "Rama:" and "X1X2:", the residue of corresponding conformer are indicated by (-), (+) and (X), which mean favored, reraly found (>5.0%) and very rarely found (>1.0%) in the statistic library.

If the program finds conformers revealing splitted clusters with largely different dihedral angles, the major (.) and minor (m) rotamers are also indicated in the child listbox.   Double-clicking the item telling you the splitted one opens another window showing the detail of the splitted angle.   The subroutine is designed to calculate dihedral angle cluster based on "Centrion clustering method".

Double-click to open the window

Structure assessment tools on ACS module
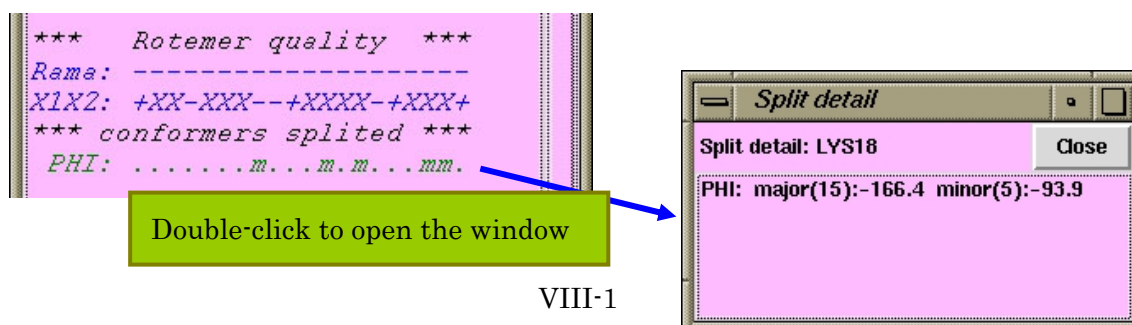
**2) The new sequence board on the magenta module**

By pressing the button of desired atom label (N, A, B, G, D, E and Z, as shown in left panel, yellow box), user is allowed to paint buttons for the assigned and unassigned atoms with different colors. The labels on the buttons, N, A, B, G, D, E, Z, and H are corresponding to the assignment types of "^{1}H-^{15}N", "^{1}H-^{13}Cα", "^{1}H-^{13}Cβ", "^{1}H-^{13}Cγ", "^{1}H-^{13}Cδ", "^{1}H-^{13}Cε", "^{1}H-^{13}Cζ" and "^{1}H-^{13}Cη", respectively.

In the bottom frame of the sequence board, there are 20 buttons (right panel, blue box) indicating **amino acid composition** of the target protein. If user click one of the buttons, user can emphasize

the selected amino-acid type in the sequence board.

The buttons found on the right-side of the sequence board, "2nd" and "ASA" are used for structure
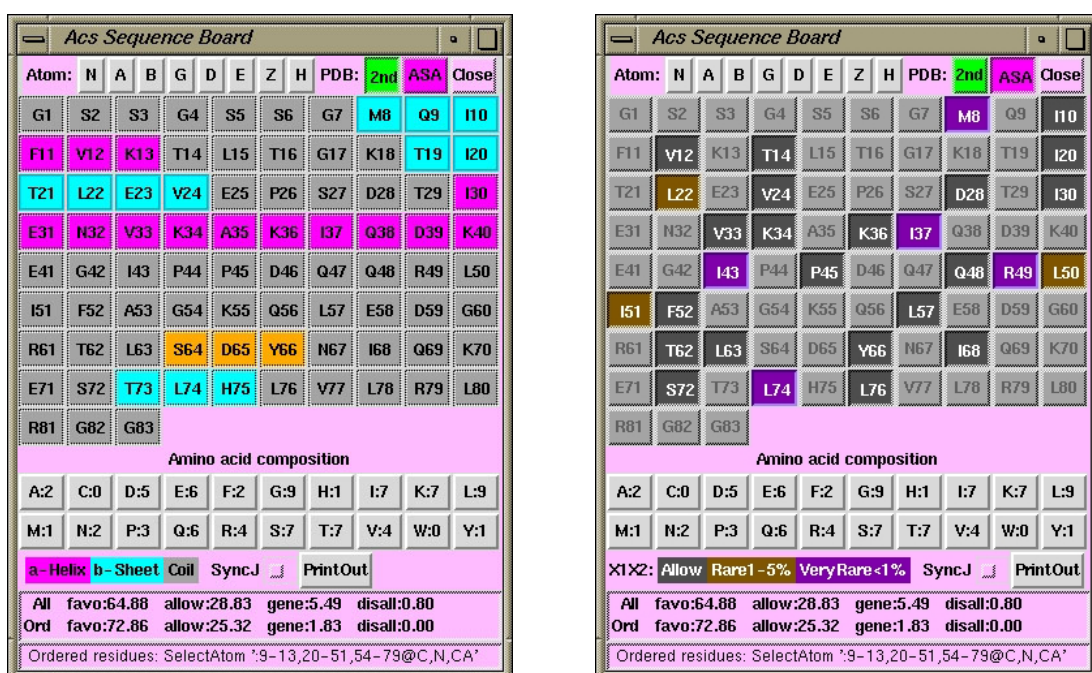
validation ased on the calculated series of structures.

Structure assessment tools on ACS module

In the secondary structure analysis mode, the detected secondary structure types are colored by;

α-Helix	magenta
3-10 Helix	orange
π-Helix	red
β-Sheet	cyan
β-Burge	light-green

In the ASA analysis mode, the ASA value of side-chain less than 30% (buried) is indicated by sunken relief button., while the others (exposed to solvent) by raised relief button.   The rotamer probability of only buried residues are calculated with the statistic rotamer library, which are indicated by different colors;

Allow (<95%)	dark gray
rarely found (1-5%)	dark orange
very rarely found (>1.0%)	dark red

3) The mechanism for detection of secondary structures

Kabsch and Sander methods [1] are widely used for secondary structure assignment tools, such as MolMol and Procheck.   The algorithm of the method is composed of two steps of simple search;

Firstly, search for hydrogen bond between α-HN (i) and α-CO (j) and classify the specific pattern of 2nd structure;

residue (i) >C=O-----H–N< residue (j)

$$E_{hbond} = f * q_1 * q_2 * (r_{ON}^{-1} + r_{CH}^{-1} - r_{OH}^{-1} - r_{CN}^{-1})$$

where r_{xy} distance between x and y atom, q_1=0.42, q_2=0.20 and f=332.   If the energy is less than -0.5 (kcal/mol), h-bond will be recognized.

Then the 2nd structure will be determined by classification of the found hbond pattern to possible 2nd structure.

n-turn(i) = HBond(i,i+n), n=3,4,5
ex) α-Helix(i,i+3) = [4-tune(i-1) and 4-turn(i)]

The new 2nd structure assignment function based on the theory has been implemented in Assign_robot.exe.
[1] Kabsch and Sander, Biopolymers. (1983) 22 (12), 2577-637.

4) Calculation of ASA (solvent Accessible Surface Area) in Kujira

The ASA values are calculated by C-program based on the algorithm established by Lee and Richards, (Lee B, Richards FM J. Mol. Biol. 1971, 55:379-400.   The results of ASA calculation by assign_robot at 0.05Å step resolution (red) perfectly matched with GETAREA (N. Özgün et al., http//www.scsb.utmb.edu/).

Structure assessment tools on ACS module
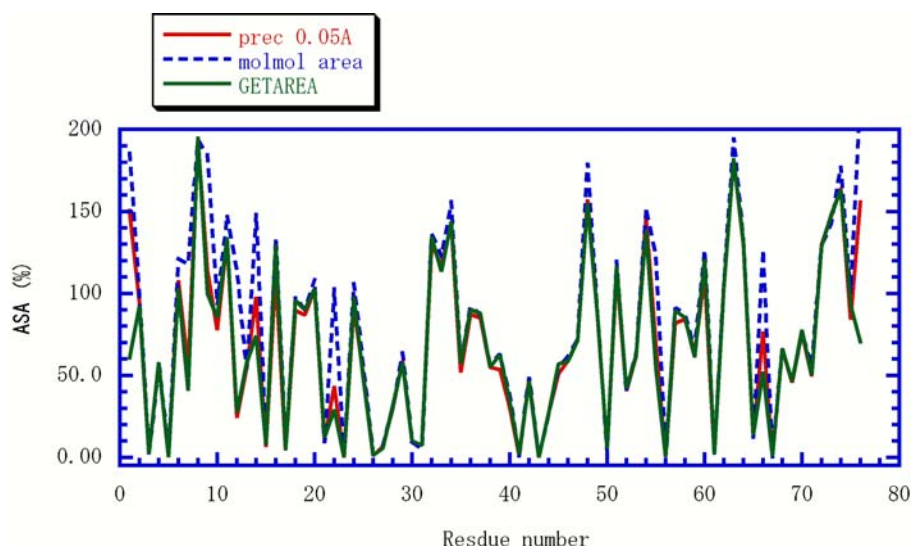
5) Command line to execute the structure validation tool

      [Syntax]

      id_robot.exe 42 <PDB file> <Ramachandran lib> <option>

      option 0: quick mode SASA   1:fine mode SASA   3:without SASA

ex) kujira/kujira_modules/bin/linux/assign_robot.exe 42 final.pdb kujira/kujira_modules/plugins/Rama_all.txt 0

6) Command line to execute automated superimposition of protein coordinates by least RMSD fitting.

The C-program implemented in Kujira calculates rotation matrix by using the conventional least RMSD fitting method (Kabsch, 1976).

The schemes for fully-automated superimposition of structures are;

        1. Determination of ordered region by order parameter calculation of phi-psi angles

        2. Cα based least root mean square fitting of structures to the first structure by quaternion method..

        3. calculate Cα rms values and search for another ordered region

        4. N, Cα, CO fitting by quaternion method to give superimposed structures.

The subroutine supports MolMol-type PDB, the authentic (the latest version of IUPAC) PDB format, CYANA1 and CYANA2 format for input and output formats.　　The current version does not supports the structure file containing hetero molecules such as Zn and nucleotides.

      [Syntax]

      id_robot.exe 46 <input PDB> <opt put PDB> <option>

      options   molmol: molmol format   pdb: IUPAC-PDB (not RCSB-PDB)

ex) kujira/kujira_modules/bin/linux/assign_robot.exe 46 final.pdb test.pdb molmol