TD ASCOM/NUS
BEST/NUS workshop 14 october 2015
Ewen Lescop and Nelly Morellet
Biologie et Chimie Structurales- ICSN Gif-sur-Yvette

Here we will process and analyze a series of NMR spectra. The nmrdata and the CCPNMR projects are located in the folder '/home/nmruser/TD_Workshop/'.

I.   Analysis of folded spectra (ASCOM)

Here we propose different efficient methods to analyze spectra collected with reduced spectral width in the ^{15}N dimension. The protein is ubiquitin and one ^{15}N HSQC spectrum has been collected with a spectral width of 31ppm centered at 115.5ppm. Then the ASCOM procedure was applied and a new spectral width of 22ppm was chosen and set for the 3D HNCO.

The two spectra are loaded in the 'ubi_ASCOM' ccpnmr project.

(1) Move to the ASCOM folder ('cd ASCOM') and load the project 'analysis ubi_ASCOM'

(2) Look at the ^{15}N HSQC in window 2. Do recognize folded peaks?

(3) Now plot the 3D HNCO spectrum by clicking on the button 'Expt_2:hnco'. This shows the projection of the 3D HNCO along the ^{13}C dimension so that all the peaks are visible. Note that due to the smaller spectral width in the ^{15}N dimension, several peaks are folded. Note also that the folded peaks are obtained by a translation of the original peaks (at the original frequencies) by the same value. Check that this shift has the value of the small spectral width (22ppm).

(4) In order to facilitate the analysis, it is convenient to duplicate the HNCO in the 15N dimension to fill the missing ^{15}N frequency range. To that aim, go to >Experiment>Spectra, select 'Expt_2:hnco', click on 'referencing' and look at the rightmost cases. Change now the 'minimum aliased frequency' and 'maximum aliased frequency' to 100ppm and 140 ppm respectively for the ^{15}N dimension. Observe the result on the window. Note that all original ^{15}N HSQC peaks are now on the top of a HNCO peak. Note also that the HNCO crosspeaks are now also duplicated but with distinct appearance: some with continuous lines, other with dashed lines.

(5) Let's now focus on getting ^{13}C information for a subset of residues.

(a) First let's focus on residue E64 on the top left part of the HSQC spectrum (9.3ppm/115ppm). In order to extract the strip from the HNCO, select the peak (it should appear with a box now), click right, then 'navigate' and select "1H 15N 13C window 3". This operation automatically extracts the 1H/13C plane at the relevant 15N frequency. The relevant frequency is the 15N chemical shift associated to the HSQC peak of E64.

(b) Let's focus now on residues I13 (9.5ppm/127.8ppm). Do the same operation as for E64. Note that although this peak is folded in the HNCO, ccpnmr easily extract the relevant plane, ie the one at 127.8-22=105.8ppm.

(6) One residue is particular: A46. In this HSQC spectrum, its cross-peak is found at 15N = 101ppm. This is outside of the ^{15}N window in the 3D HNCO and is expected to be at 101+22=123ppm. However no peak is present at this position and the navigation in the 3D points to a region devoid of signal.

The inspection of the HNCO suggests that the HNCO contains a peak at 101ppm superposed with the HSQC peak of A46. However, note that are not exactly superimposed, so they can correspond to the same residue. Propose an explanation for all these observations.

Since this residue has never been observed with its "real" ^{15}N chemical shift, it might be a problem. Actually, CCPNMR can handle those tricky situations. The absolute frequency of A46 is set to 101ppm. This information can be retrieved from the following: select the peak, type 'a'. This idea is to change the absolute frequency of A46 by unaliasing its frequency. This is done as follows: select the peak, click right on the peak, and, go to Peak>Unalias Peak and change the value for "Num. aliasing" to ±1, ±2,…. This will automatically change the ^{15}N absolute chemical shifts by adding or subtracting a multiple number of times the current spectral width. Set the value to -1 and note that the A46 HSQC cross-peak within the original HSQC spectral window changed to dashed line while the duplicated HSQC region (15N higher than 131ppm) contain the A46 with continuous line. This is the indication that a peak is at its real apparent frequency and not folded (this is not indication that it is the physical value). Let's now select the "unfolded" A46 cross-peak (the one at 133ppm) and navigate to the window 3. Note that there is now some signal here. Similarly, the projection of the 3D HNCO window 2 also displays a peak corresponding to A46. Conclude that although none of the two spectra measured the absolute ^{15}N chemical shift of A46 because it was folded in both spectra, the tools in CCPNMR allow to easily detect this situation and handle it.

(7) For the correct analysis of all 3D spectra, the "Num. Aliasing" should be consistently and carefully done. Indeed, the final chemical shift value of a given atom (resonance) is averaged out over the position of all peaks to which this resonance is assigned out. Any missetting of the "Num. Aliasing" will created large mistake. Such situation can be easily detected as follows: Go to Resonances>Resonances. The shift and SD columns show the averaged and standard deviation chemical shift values, respectively. Missetting can be detected by too large SD values. Clicking on "SD" will order the resonances according to the SD values, thus highlighting possible outliers. Fortunately, when one uses the "Pick and assign from Roots" tool, this is automatically done. Therefore as soon as the HSQC peak list is correct (with folded peaks such as A46 identified), the effect of aliasing is transparent to the operator when he/she is analyzing the 3D spectra.

Close the project: Project >Close

II.   Processing of 2D spectra by qMDD : the case of strychnine

The NUS principle can be applied to a variety of NMR pulse-sequences and sample. Here we will illustrate on strychnine the efficiency of NUS for homo- and heteronuclear pulsesequence of various signal compositions.
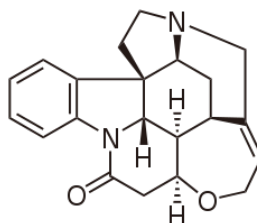
Structure of strychnine

Series 10–14	Series 20–24	Series 40–44	Series 60–64	Series 70–74
TOCSY	HSQC	COSY DQF	NOESY	HMQC

For each pulse-sequences, we collected a series of experiments with various sampling schemes. They are in the folder '/home/nmruser/TD_Workshop/strychnine_NUS_04092015'. For each pulse-sequence, one series starts with the number **x0** (the number depends on the pulse-sequence) which was collected with the **conventional approach** providing a long experiment with moderate resolution, albeit easy to process using regular Fourier transform. The experiment **x1** was collected using the NUS principle with **100% of points**. It is therefore identical in total time and resolution compare to x0, and can be used to verify the processing procedure. The experiment **x2** was collected in the same experimental time as x0 and x1 using **25%** points and TD multiplied by 4 compared to x0/x1, resulting in improved resolution in the same total time. The experiment **x3** was collected using **25%** points and the same TD as compared to x0/x1, resulting in similar resolution in total time divided by four. The goal of the work here will be to be familiar with the processing tools and to assess the effect of NUS processing on the final spectra. All spectra have already been processed and loaded into the project 'nus'.

A.   Comparison of the 13C HSQC spectrum collected in conventional manner and using NUS

Here, we will inspect the original ^{13}C HSQC spectrum collected and processed in the conventional manner (folder 20), learn how to process the NUS spectrum collected in folder 21 with qMDD and use these tools to process folders 22 and 23.

qMDD uses multi-dimensional decomposition (MDD) and compressed sensing (CS) to replenishing missing data points in the full matrix followed by regular FT processing of the complete data. The software offers three general possibilities (i) direct Fourier transform, all missing points are set to zero (Method Zero or FT); (ii) multi-dimensional decomposition (MDD); and (iii) compressed sensing (CS).

qMDD software used nmrpipe for spectral data conversion and traditional processing. Two scripts used for that:
*-**fidSP.com** is responsible for conversion of the spectrum from Bruker format (file* ser*) to the nmrPipe format and for processing the directly detected dimension. It may be modified, for instance, for adjusting phase in the indirect dimension or chemical shift references.*
*-**recFT.com** processes all indirect spectral dimensions after the missing data in the time domain interferogram is replenished by mddNMR. The scripts can be viewed and edited from "Advanced" display by pressing the corresponding buttons.*

*The calculations are performed by the master script **proc.sh**. The script can be viewed and modified in "Advanced" view by pressing button proc.sh or in any text editor. It can be executed either by clicking on the 'Run' button in qMDD window or by typing '/proc.sh' in the terminal. The file is as follows and contains all parameters that are also visible in the qMDD GUI:*

#!/bin/tcsh
setenv FID ../23

```
setenv fidSP fidSP.com
setenv REC2FT recFT.com
setenv in_file nls.in
setenv selection_file nuslist
setenv FST_PNT_PPM          8.5
setenv ROISW                8.5
setenv proc_out test.dat
setenv SPARSE          y
setenv NUS_TABLE_ORDER      '1'
setenv NUS_TABLE_OFFSET     0
setenv NUS_POINTS                25
setenv NI                        '25 1 '
setenv NIMAX                '100 1 '
setenv NDIM                 2
setenv MDDTHREADS           2
setenv METHOD                    FT
setenv CS_alg               IRLS
setenv CS_norm              0
setenv CS_lambda            1.0
 setenv CS_niter            10
setenv CS_VE                y
setenv SRSIZE               0.1
setenv NCOMP                10
setenv NITER                200
setenv MDD_NOISE                 0.7
setenv CEXP                 nn
setenv CT_SP                nn
setenv lambda               0.001
mddnmr4pipeN.sh  1 2 3 4 5
set ecode=$? ;if ( $ecode ) exit($ecode)
```

The processing is performed by mddnmr4pipeN.sh *command at the end of the script.* *As arguments, this command takes list of tasks to be done, which are defined by unique numbers.*

0   – print full list of parameters recognized by the program.
1   – conversion of ser/fid to nmrPipe format; processing of the direct dimension and extraction of region of interest (ROI) (see parameter fidSP).
2   – preparing input for MDD calculations.
3   – MDD calculations over all sub-regions of the ROI.
4   – full reconstruction is produced from MDD components and residuals (see MDD_NOISE).
5   –the full time domain reconstruction obtained in step 4 is processed using an nmrPipe script (see parameter REC2FT)

The original HSQC in folder "TD_Workshop/strychnine_NUS_04092015/20" was processed using nmrPipe (without qMDD). The experiment was loaded together with all other processed spectra in the ccpnmr project (in a terminal window go to "TD_Workshop/strychnine_NUS_04092015/ and type "**analysis nus**"), look at window HMQC/HSQC and more precisely at the HSQC1_ref spectrum. Note that there are few overlaps in the ^{1}H dimension. Therefore for most experiments their will be only few peaks in the indirect dimension. This is typical of NMR spectra of small molecules and largely different from the crowded protein NMR spectra (see next chapter)

B.   Processing of the spectra collected with non-uniform or non-linear sampling (*NUS*)

1.   To process the experiment 21 (the experiment was collected using 100%)

In a terminal go to "TD_Workshop/strychnine_NUS_04092015/21" start qMDD (type '**qMDD**') in the terminal. A window is opened which invites to select a spectrum for processing. Click on "**open spectrum**" (A window is opened) (the directory is chosen automatically) and click on "**choose**" (a new window is opened) and a question is asked: *21.proc dir already exists, overwrite*, you answer "**no**" or "**yes**" to this question. Indeed the 21.proc directory may exist from previous calculations. If you answer 'no' you will use the already parameters adjusted for this experiment. If you select 'yes' this will create or overwrite the existing folder 21.proc containing the automatically generated files with default parameters. Here we will start from scratch and select 'yes' to see how to modify the relevant parameters. In the qMDD window you can see that the Proc.dir is now define in "**Proc.dir**" tab.

We will use direct Fourier transformation, since all the points were collected, activate on the qMDD window "Data conversion", "**Zero**", "Fourier transformation". You are ready to process the spectrum by pressing the button "**RUN**" and wait until output in the terminal window indicates successful completion of the calculations. Look at the spectrum the result click on "**nmrDraw**" on the qMDD window. The nmrDraw window is open, to see the spectrum, click right on the mouse on "**File> select file**" test.dat and "**Read/Draw**" (bottom of the window). What do you think of the calculated spectrum ?

As you can see the ***spectrum has to be phased***. It can be done with nmrDraw. But here you have just to insert the right value in the **fidSP.com**. Open the fidSP.com on the qMDD window. The phase parameters can be written as shown in red below.

```
bruk2pipe -in ./ser -bad 0.0 -noswap -DMX -decim 2376 -dspfvs 20 -grpdly 67.9872589111328\
-xN          2560          -yN     50                         \
-xT          1280          -yT     25                         \
-xMODE       Complex       -yMODE       Echo-AntiEcho   \
-xSW         8417.508      -ySW         28089.888       \
-xOBS        699.550       -yOBS        175.902         \
-xCAR        4.711         -yCAR        70.010          \
-xP0         -17.0         -yP0         90.0            \
-xP1         -3.4          -yP1         0.0             \
-xLAB        1H            -yLAB        13C             \
```

An alternative to this method is to modify the parameters in the fidSP.com and recFT.com files as:

```
-ndim 2  -aq2D States \
| nmrPipe -fn POLY -time                              \
| nmrPipe -fn SP -off 0.48 -end 0.95 -pow 2 -c 0.5  \
| nmrPipe -fn ZF -auto                          \
| nmrPipe -fn FT -auto                          \
| nmrPipe -fn PS -hdr                           \
| nmrPipe -fn PS -p0 -17.0  -p1 -3.4  -di                \
| nmrPipe -fn POLY -auto -xn 5.0ppm -ord 1        \
| nmrPipe -fn EXT -sw                    \
| pipe2xyz -z -out ft/data%03d.DAT -ov -nofs -verb
```

And in **recFT.com**

```
| nmrPipe  -fn TP -auto                              \
| nmrPipe  -fn SP -off 0.48 -end 0.95 -pow 2 -c 0.5  \
| nmrPipe  -fn ZF -auto                              \
| nmrPipe  -fn FT                                    \
| nmrPipe  -fn PS -hdr                               \
| nmrPipe  -fn PS -p0 90.0 -p1 -0.0 -di              \
| nmrPipe  -fn TP  -auto                             \
```

You must adjust the MddNMR parameters corresponding to region of the spectrum that has to be calculated. You can try with different values and see the results. With "**first point ppm**" set to **9** and "**region of interest SW**" set to **8.5** and click on "**save**" you will calculated all the spectra (ie between 9ppm and 9-8.5=0.5ppm). In a second time try with "**first point ppm**" set to **4** and "**region of interest SW**" set to **2** (the [2-4]ppm region) and click on "**save**". Sometimes it is useful to first decrease the spectral width for the calculation in order to also reduce the calculation time proportionally to the region size, to adjust the various parameters, and in a second step, when all processing parameters are correctly set, the full region can be calculated. With the new phase parameters, process again the spectrum ("Run") and update the display in nmrDraw with the newly processed experiment by left-clicking on "File" in nmrDraw. Check that the processing is ok now.

2.   To process the experiment 23 (the experiment was collected using 25% points and the total acquisition time divided by four)

In a terminal go to "TD_Workshop/strychnine_NUS_04092015/23" start qMDD (type '**qMDD**') in the terminal. A window is opened which invites to select a spectrum for processing.

Click on "**open spectrum**" (A window is opened) (the directory is chosen automatically) and click on "**choose**" (a new window is opened) and a question is asked: *23.proc dir already exists, overwrite*. Answer "**no**" to this question (you will use the already Fourier transformation parameters adjusted for this experiment, ie mostly the same as for the experiment 22).

a)   Using direct Fourier transform, all missing points are set to zero:
Activate on the qMDD window "Data conversion", "**Zero**", "Fourier transformation".

You are ready to process the spectrum by pressing the button "**RUN**". Look at the result on "**nmrDraw**" on the qMDD window. The nmrDraw window is open, to see the spectrum, click right on the mouse on "**File> select file**" test.dat and "**Read/Draw**" (bottom of the window). The spectrum appeared, are you satisfied with what you see? It provides poorer results due to massive aliasing artefacts. Nevertheless, "**FT/Zero**" mode is useful since it allows fast look at the spectrum and adjustment of nmrPipe processing parameters (e.g. phases).

b)   Using multi-dimensional decomposition (MDD)
Activate on the qMDD window "Data conversion", "**MDD**", "Fourier transformation".

Prior to "RUN", you may define a small region of interest by setting "First point ppm" and "Region of Interest SW" (with "First point" as its downfield border) in ppm followed by pressing "save" button. This reduces the calculation time proportionally to the region size.

You are ready to process the spectrum by pressing the button. A new terminal window opens in which script script *proc.sh* is executed "**RUN**" and wait until output in the terminal window indicates successful completion of the calculation. You can open this file

(click on proc.sh) to get an idea of its composition and recognize some of the parameters. Again look at the results on "**nmrDraw**" on the qMDD window. The nmrDraw window is open, to see the spectrum, click right on the mouse on "**File> select file**" test.dat and "**Read/Draw**" (bottom of the window). The spectrum appeared, are you satisfied with what you see?

qMDD allows modification of several most important parameters and scripts. This is done in the "advanced parameters"

#threads: Default value is 2, the maximal number of parallel processes
NCOMP: Number of MDD components per sub-region (default value 30)
It correspond to the number of peaks (and respective line shape/position) used to fit the available data within a subregion/strip. This parameter affects quality of the solution. In most cases, a default value of 30 for a sub-region strip of 0.1-0.2 ppm (parameter SRSIZE) in the directly detected ^{1}H dimension is a good guess. The number of components must be 20-50% larger than number of expected cross-peaks for 2D's and triple resonance backbone experiments, or number of diagonal peaks for 3-4D NOESY/TOCSY experiments.

RMDD: CEXP and CT_SP flags can be used in any combination for the dimensions:

CEXP: y/n toggle R-MDD / MDD mode for a dimension, with 'y' time domain shape in the dimension is expected to be autoregressive, assume that the FID in the dimension is a complex exponent. CEXP=y may be used, for example, for HNCO and HNcoCA experiments, but not for the NOESY's.
CT_SP: n/y toggles mirror image processing for dimensions with CEXP='y'. Set 'n' for the first indirect dimension. CT_SP have to be set to 'y' for the constant-time second indirect dimension (typically 15N) in the triple resonance experiments.

NOISE: A factor that scales residuals of the MDD calculations as they are added to the reconstructed spectrum: 0 -1
NTER: Number of iterations for MDD (50 to 500)

(1)   To show the effect of **NTER**

Try again by changing for example NTER 200, compared with NTER 50. As you can see increasing the number of iterations, decreasing the noise streaks in the spectrum but the calculation takes longer. Try again with NTER 500.

(2)   To show the effect of **NCOMP**
Compare the spectra calculated with NTER 200 NCOMP 25, and NTER 200 NCOMP 10. This parameter seems important.

(3)   To show the effect of **NOISE**
Compare the spectra calculated with NTER 200, NCOMP 10, NOISE 0 and NTER 200, NCOMP 10, NOISE 1 (no significant differences)

c)   Using compressed sensing (CS).
Activate on the qMDD window "Data conversion", "**CS**", "Fourier transformation". Apply the protocol already described above. What do you say about the spectrum now ?

Modifications of several most important parameters and scripts. This is done in the "advanced parameters"
CS_niter : Number of iterations for CS (default value 10 is ok for IRLS). Change to 100-10000 for IST.

CS_alg: CS algorithm: IRLS –iterative reweighted least squares, or IST –iterative soft thresholding (IRLS by default)

CS_lambda: CS regularization (default value 1.0 is ok for all studied cases)

CS_norm: Norm for CS IRLS algorithm: 0 -1

You can try to modify these parameters in order to evaluate their effects on the spectral quality.

d)   Determine the minimum % points to be recorded in order to obtain a correct spectrum

Here we will see how the spectral quality progressively deteriorate when reducing the NUS percentage. We will start from the experiment 21 collected with 100% NUS and keep only the first x% of the collected increments for processing. A script has been written to extract the % of NUS points we want to test: extract_firstpoints2D.csh. It is located in the TDworkshop folder. You can open it using 'j extract_firstpoints2D.csh' in the terminal. This script will automatically use the CS processing method with default parameters.

In a terminal go to TD_Workshop/strychnine_NUS_04092015/, write "extract_firstpoints2D.csh **21 5 0**" this command will extract 5% of the total points collected in experiment 21 => the directory 21.proc_extract_5 is created and the nmrDraw window is open. You can evaluate if 5% of the points is enough to recalculate a correct spectrum. Try with 10% or 50% of the points. The last option '0' corresponding to the possibility to add noise during processing (see below). Here no noise is added.

C.   Different examples of NUS data

We will rapid explore different (homo- or heteronuclear) experiments with different signal content (number, distribution). As explained before, for a given pulsequences (HSQC, HMQC, COSY, TOCSY, NOESY) a series of experiments was collected as reference (x0), 100% NUS (x1), 25% NUS but four times longer TD (ie increased resolution for same total duration, x3) and 25% NUS (same resolution but four times faster). All experiments have been loaded into the 'nus' ccpnmr project in the folder strychnine_NUS_04092015. To load it, go to the directory and type '**analysis nus**'. This opens ccpnmr and the first window 'HMQC/HSQC' that displays the four spectra collected for each of the two heteronuclear experiment (name: PP1_ref/ PP2_refNUS/PP3_NUS1/PP4_NUS3 for the x0, x1, x2 and x3 spectra respectively, PP is the pulsequence). The first visible spectra are HSQC1_ref and HSQC2_refNUS that are almost identical. This is normal because the only difference is the order by which the increment was collected. Compare HSQC2_refNUS with HSQC3_NUS1 and HSQC4_NUS4. To this end, click on 'Spectra' in window 'HMQC/HSQC' and click on the wanted names to display or remove them. Note that HSQC3_NUS1 shows increased resolution compared to HSQC2_refNUS, although they were collected in the same total duration. Similarly note that HSQC4_NUS2 is of similar quality as HSQC2_refNUS, despite the four times faster experimental time. To change the levels, use 'r' or 'e' to decrease or increase the base levels for all visible spectra in the current window.

The HMQC experiment contains more signals than HSQC. Do the similar analysis as for HSQC.

Now, display the 'TOCSY' window by in the main ccpnmr menu, to the **Window>HH TOCSY**. Note that despite the much larger number of crosspeaks, the high dynamic range (weak vs strong peaks), positive/negative signals, etc…. the CS processed

spectra are of high quality. The same observation can be made for NOESY (**Window>HH NOESY**).

The COSY crosspeaks have a complex structure and this is also conserved in NUS processed (window COSY).

Note that for all spectra, at 25% NUS, the artefacts are very weak, if visible. The signal-to-noise is also quite good despite the reduced experimental time. This demonstrates that for small molecules, NUS can be very useful used to speed up data collection while retaining high resolution, or obtaining much higher resolution within the same duration.

III.   Processing of protein 2D/3D/4D NMR spectra: ubiquitin

Here a series of spectra were collected on a 1.2mM ^{15}N-^{13}C labeled ubiquitin. They are located in the folder ubi_BEST_NUS_092015. The experiments numbered 30 to 38 and 40 to 48 correspond to the following pulsequences (BEST/TROSY):

30/40	31/41	32/42	33/43	34/44	35/45	36/46	37/47	38/48
BTROSY	BT_HNCO	BT_HNCACO	BT_HNCA	BT_HNCOCA	BT_HNCOCACB	BT_HNCACB	BT_HNCOCANH	BT_HNCANH

The 30-38 series was collected as regular 3D spectra with 100% NUS. They will be used to learn how to use qMDD to process 2D/3D spectra, to estimate the effect of reducing the NUS percentage, as well as estimating the effect of low or high signal-to-noise ratio on NUS. The total experimental time was 13h30. The second series was collected using 10% NUS resulting in a total experimental time of 1h30.

A.   Processing of 2D ^{15}N BEST-TROSY

Let's first have a look to the 2D ^{15}N TROSY spectra (expno 30). We will use qMDD to process it.

1. Start qMDD by typing 'qMDD' and open the folder 30. It creates the 30.proc with preset parameters. We will first use the FT processing mode since all points have been collected (100% NUS). This can be used to help adjusting the processing parameters. Compared to the default parameters, we will change the following parameters:

a. Change the "first point ppm" to 9.6 and the "region of interest SW" to 3 and click on 'Save', in order to focus on the region of the spectrum containing signal. This speeds up signal processing.

b. Click on '**Run**' and then on '**nmrDraw**'. The phase is not correctly set. Click on 'fidSP.com' and change the following lines to set the 1H (xP0/xP1) and 15N (yP0/yP1) phase:

-xP0   36.8   -yP0   24.0   \
-xP1   46   -yP1   0.0   \

c. Click again on '**Run**' and in '**nmrDraw**' click on 'File' to update the display with the newly processed spectrum. Note the quality of the spectrum.

2. We will now progressively reduce the number of points used for processing. We collected initially the full matrix (TD from 1 to TDmax) in a random manner. Here we will keep the y% first points which mimics the situation were y% of points were really collected. This allows to assess the effect of MDD/CS processing when less and less points are acquired. The script extract_firstpoints2D.csh has been written to keep the y% first points and adjusts all qMDD parameters. It is used as follows: "**extract_firstpoints2D.csh X Y Z**" where X is the experiment number, Y the percentage and Z the added noise. It copies the folder x.proc to x.proc_extract_y and adjust the parameters in folder x.proc_extract_y. It also proposed to add

noise to the spectra (Z). The script set the processing method to CS and its default parameters, whatever the original method is.

We will first keep only 50% of the points from the experiment. So we type "**extract_firstpoints2D.csh 30 50 0**". This automatically does the processing and open nmrDraw to display the resulting spectrum. Compare the spectrum with the reference one by opening nmrDraw in the 30.proc folder. Do no notice a any difference? Now use only 25, 10 then 5%. Note that the quality of the spectrum progressively decreases. While the artefacts are already visible from 25%, they are really annoying at 10% and even more at 5%. Many artefacts can be removed by increasing the number of iterations for CS. To do that, move to the folder 30.proc_extract_10, open the proc.sh file (type '**j proc.sh**'), change the CS_niter to 100, save and close the file, and finally execute '**proc.sh**' in the terminal by typing '**proc.sh**', finally update nmrDraw display by clicking on the 'file' button). Compare with the initial processing (10 iteration).

3. The ubiquitin sample gives very nice spectra but what about samples with **less signal-to-noise ratio**. Here we will add noise to the original spectrum (using the addNoise from nmrPipe, see script extract_firstpoints2D.csh). Execute the command '**./extract_firstpoints2D.csh 30 100 300000**' to see the level of added noise to the 100% spectrum. You can extract a row by typing '**h**' in nmrDraw and move the cursor. Leave this mode by typing '**e**'. Now reduce the percentage to 50 ('./extract_firstpoints2D.csh 30 50 300000). Extract the row at the same position (peak) in the two spectra to assess their respective signal-to-noise ratios. Go down progressively to 10% and 5%. Note that at 5% the number of artefacts is detrimental. At 10%, most peaks are visible. Observe their peak lineshape. Observe the structure of noise and compare to 100%.

B.   Processing of 3D ^{15}N BEST-TROSY HNCO

1. The 3D HNCO 100% NUS is available in number 31. Start qMDD and open the experiment. In order to focus on the 1H spectral region with "first point ppm" and "region of interest SW" set to 9.6 and 3 respectively. Click on '**Run**' (it takes more time than 2D) and open '**nmrDraw**'. Right-click on File, then on '**select File**' and select '**1H.15N.dat**' and click on '**Read/Draw**'. This corresponds to the 2D projection of the 3D HNCO spectrum along the 13C dimension. This allows to have a general overview of the 3D. Clearly the phase is not correct along the 1H dimension. Have a look to the '**1H.13C.dat**' to check the 13C dimension. It seems ok. Now set the 1H phase (-xP0 224.2 –xP1 36) and run again the processing. It is clear now that the 15N phase can be improved. Correct it by setting (-zP0 107.5 /-zP1 -7.0) in the file fidSP.com and process again. Once you are satisfied with the projection you can explore the 3D spectrum (select the file **31.ft3**). Move along the planes by changing the number **z** (upper left corner). A polynomial correction could be applied along the 1H dimension by inserting the line '**| nmrPipe –fn POLY -auto**\' just after that containing 'ZTP' (In the 31.proc directory open the recFT.com file '**j recFT.com**'). Run again the processing and check the result.

2. We will now explore the effect of reducing the percentage. The script extract_firstpoints3D.csh does the job (**'./extract_firstpoints3D.csh 31 5 0'**) for 5%. Inspect the ^{1}H-^{15}N and ^{1}H-^{13}C. Then observe the 3D spectrum (31.ft3). Note that artefacts are visible in the 3D but that they cancel out in the projections. However note the artefacts are weak compared to real signals.

3. We will now check the effect of noisy spectra. We will add noise the 100% spectrum by executing './extract_firstpoints3D.csh 31 100 2000000'. Unfortunally, the CS algorithm takes too much time. We will instead use FT. So, stop the process by typing 'Crtl+C'. Move to the folder by typing '**cd 31.proc_extract_100**' and edit the proc.sh file ('**j**

proc.sh'). Change the METHOD from CS to FT and save. Execute '**proc.sh**'. This time it is much faster. Open nmrDraw ('**nmrDraw&**') and open the '1H.15N.dat' spectrum. Now reduce the percentage to 5%. Note that the signal-to-noise dropped. Go now to 15% and compare.

C.   Analysis of series of triple resonance experiments.

The whole series of 3D triple resonance experiments (HNCO, HN(CA)CO, HN(CO)CA, HNCA, HNCOCACB, HNCACB) plus the HN(CO)CANH and HN(CA)NH was collected twice : 100% NUS within 13h and at 10% NUS within 1h30. They were all collected using the BEST-TROSY principle. They are all loaded in the ccpnmr project 'ubi_workshop'. Load it by typing '**analysis ubi_workshop**'.

The window HSQC contains the BEST-TROSY collected at 100% (spectrum 'trosy') or 50% (spectrum 'trosy_nus50'). Compare them in terms of artefacts, spectral quality, signal-to-noise ratio. You can display the assigned peaklis of the 'trosy' spectrum by clicking on '**Peaks**' in window HSQC and then on 'Expt2:trosy:1'.

Then go to the window 'HNCO'. It contains the HNCO and HN(CA)CO spectra collected either as 100% or as 10% (with the _nus10 extension). Several strips are shown for consecutive residues (we know the assignment). They were obtained using the following steps:

1. in the main ccpnmr menu, go to Resonance>Spin system

2. Click on 'Residue' to order the column according to the residue number. Select the range '2Gln' to '9Thr' by selecting first '2Gln', and the Shift+click on '9Thr'.

3. In the upper right region, select the window 'HNCO' and click on 'Display Strips'.

By adjusting the display levels of visible spectra in HNCO window with the 'r'/'e' shortcut you can estimate the difference in signal-to-noise ratio when collected only 10% of data points. Because here the SNR is very high the NUS reconstruction is very efficient and restore the total SNR even if only 10% signal is really collected.

Now open the 'HNC' window. It contains the 100% HNCOCA, HNCA, HNCOCACB, HNCACB. You can then observe in the 'HNC_nus' window the same spectra collected either in 100% and 10%. Imagine that you have to guess which spectrum was collected with or without NUS, would you be successful?

For relatively small proteins or unfolded proteins, the HNCANH and HNCOCANH experiments can be extremely useful to connect the 15N atoms from consecutive residues. They are shown in window 'HNN' with or without NUS.

We also collected a 4D spectrum: ^{15}N-HSQC-NOESY-^{13}C-HSQC. The parameters were:

	^1H (direct)	^{15}N	^1H (indirect)	^{13}C
SW	16	21 (ASCOM)	8.0123	77.66
TD	2k	32	64	32

We use NS=8, 7%NUS. Total duration: 12h.

Go to window 4D_HC and inspect the planes by modify the 1H/15N chemical shifts with the cursor at the bottom of the window. Now go to the HSQC window and select a cross peak, for example for residue 64. Right click and go to **Navigate>'- - 1H15N'** in 4D_HC. The 1H-13C plane displayed in window 4D_HC contains 1H/13C correlation peaks for all 1H that give a NOE crosspeak to 1HN of residue 64.

TOOLS when using Virtualbox

To copy and write in the terminal: ctrl+shift+c anc ctrl+shift+v???