**Instructions for including ncAAs in CCPNMR v3**

Anne Conibear (based on CCPNMRv3 documentation, forum posts and some trial and error)

* Assumes you have an unnatural or modified residue (ncAA) in a peptide/protein chain with a standard backbone.
* Assumes spectra are loaded in CCPNMR and you are familiar with basic operations in CCPv3 (https://www.ccpn.ac.uk/v3-software/tutorials).

**Including ncAAs for assignment in CCPNMRv3**

**Case 1**: If the ncAA contains additional atoms/resonances but they don’t show up in your NMR spectra (e.g. sulfate group on sulfotyrosine, phosphate group on phosphoserine if using H/C/N spectra):

* include the residue as the closest canonical residue (e.g. tyrosine, serine, respectively) in the Chain, generate an NmrChain and assign as normal (https://www.ccpn.ac.uk/v3-software/tutorials/backbone-assignment-tutorial/view).
* export .nef file and modify for CYANA (see below).
* download or generate a CYANA library (.lib) file and append to standard library file.
* make sure residue names (e.g. TYS for sulfotyrosine) match the name of the residue in the .lib file.
* generate .seq, .peaks, and .prot files from .nef file (see below). Run CYANA.

**Case 2**: If the ncAA contains atoms/resonances that you want to assign and include in structure calculations (e.g. acetyllysine, methyllysine, methylarginine in which the -CH3 group might have NOE peaks to other atoms), based on <https://www.ccpn.ac.uk/v3-software/documentation/faq#ChemBuildXml>:

Worked example for Ac-Gly-Gly-Arg(Me)-Gly-Gly-NH2

1. Follow instructions as for CCP v2 to build your ncAAs in Pymol and submit to the ATB. Download the output .pdb files (with CONNECT information) and .lib from the CYANA tab in the ATB.

2. Convert the .pdb file to a ChemComp file using ChemBuild (incorporated in CCPv3).

- Open ChemBuild (ChemBuild.bat in CCP directory).

- Import>PDB file>(pdb file with CONNECT information) e.g. VJI9\_760861\_Connect.pdb for Arg(Me) -> includes pseudoatoms but no connect information for these atoms.

- Check that the molecule looks sensible and atom names are correct (should be the same as the .lib file).

- Export> CCPN ChemComp XML File -> Give the residue a 3-letter code (EJG used for Arg(Me) for consistency with v2 example). No file extension > Save -> a folder with name EJG is generated in your chosen directory. Within this folder is the .xml file with a long and complex name -> \*don’t change this!\*

- Do the same process for GCE and GN2.

3. Import the molecule(s) into CCPNv3:

- Set up your CCPNv3 project and load the spectra. Create your molecule chain using the closest canonical residues (or alanine) in place of the ncAAs. e.g. GGRGG.

Graphical user interface, text, application

Description automatically generated

* Generate an NmrChain from this Chain (this will be NC:A)
* Import your .xml file: Macro>Run CCPN Macros > ChainFromChemComp

Select the path to the .xml file within the EJG folder. Give the residue a 3-letter code (EJG) and residue number (3). This will go into chain B. Select ‘Expand Atoms from Atom Groups’ and ‘Add Non Stereo-Specific Atom’.> OK

Graphical user interface, text, application, email

Description automatically generated

The new residue will appear as a new substance (5db102ab) and a new chain (B).

* Repeat the New Chain from ChemComp process for GCE (residue 1, chain C) and GN2 (residue 5, chain D).
* Create a new NmrChain from each of chains B, C and D.
* Graphical user interface, application, Word

  Description automatically generated

This creates a new NmrChain (B) with and NmrResidue (EJG) comprising the atoms (and pseudoatoms) for Arg(Me).

4. Incorporate the ncAAs into the main NmrChain (A):

- Delete the corresponding NmrResidue ‘placeholder’ in NmrChainA e.g.NR:A.3.Arg (right click on the NmrResidue and select ‘delete’).

- Edit the ncAA NR.B.1.EJG (double click) and change it from chain A to chain B, with the appropriate sequence code. Select ‘Merge to Existing’.

Graphical user interface

Description automatically generated

The main NmrChain should now contain the ncAA at the appropriate position and if you expand it, should have all the relevant atoms and pseudoatoms for assignment.

* Repeat the process for GCE and GN2
* You can now assign the spectra.
* A picture containing text

  Description automatically generated
* The problem with this approach is that residues are not "covalently linked" in the project, so backbone assignment module would not work, however this solution is ok for assigning peaks. The next issue is that when exporting to NEF file, chain information is stored in molecular system saveframe, which has to be modified manually (unwanted alanine replaced with correct modified residue and chain B deleted). We plan to improve support for small molecules and non-standard amino acids in future releases, but hard to predict the timeframe (Luca recons that code changes would require 2 - 3 weeks of his time.)

**Exporting NEF files from CCPNMRv3**

<https://www.ccpn.ac.uk/v3-software/tutorials/howto-nef-import-export/view>

**Generating .seq, .peaks and .prot files from a NEF file for use in CYANA**

Open exported .nef file in a text editor. Change:

save\_nef\_nmr\_meta\_data -> \_nef\_nmr\_meta\_data.format\_version 0.91

save\_nef\_nmr\_spectrum\_GR2919\_TOCSY1`1` -> \_nef\_nmr\_spectrum.experiment\_type NOESY (or set experiment type in CCP before exporting)

\_nef\_spectrum\_dimension' \_loop \_nef\_spectrum\_dimension.axis\_code

1 ppm H1 749.7835083 16.67147898 13.12757051 circular true .

2 ppm H2 749.7835083 13.33718319 11.46041975 circular true .

save\_nef\_molecular\_system -> \_nef\_sequence.residue\_name The sequence is exported from the Chain so you will have to change the residue names of your ncAAs to match the residue names in your chemical shift list (e.g. ALY, TYS) and .lib file.

e.g. peptide with sulfotyrosine (TYS)

Copy TYS.lib file into the folder ‘lib’ in cyana-3.98.13

In init.cya -> add in line read lib TYS.lib append

In terminal, navigate to folder where the .nef file and other cyana files are. Open cyana > cyana make sure cyana reads your new .lib file.

> read nef CX0002\_CCPproj5.nef

> write seq CX0002.seq Then open .seq in text editor and remove the chain code (e.g. A2 -> 2)

> write prot CX0002.prot

> write peaks CX0002.peaks (make sure to select peaks and calculate volumes before exporting from CCPN)

**References and links**

CCPN YouTube explanation of Chain/Residue/ResidueType/Atom and NMRChain/NMRResidue/NMRAtom: <https://www.youtube.com/watch?v=DS9IZzNsBbQ>

Explanation of how to include new substances in CCPNv3 <https://www.ccpn.ac.uk/v3-software/documentation/faq#ChemBuildXml>