

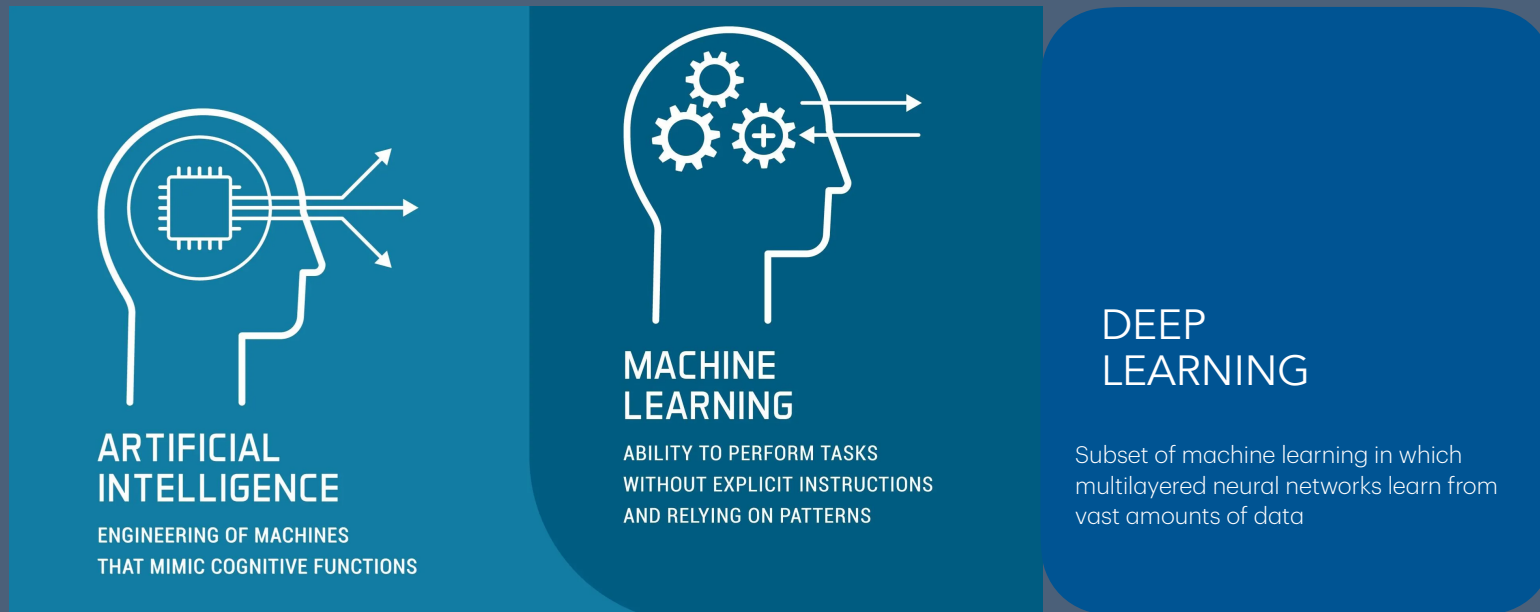
# Deep Learning Tools in NMR

Annual meeting of Managers of UK Magnetic Resonance Facilities

Angelo M Figueiredo - University College London

# Introduction

AI, machine learning, and deep learning are interchangeable and easily confusing



A program that can sense, reason, act and adapt

Algorithms whose performance improve as they are exposed to more data over time

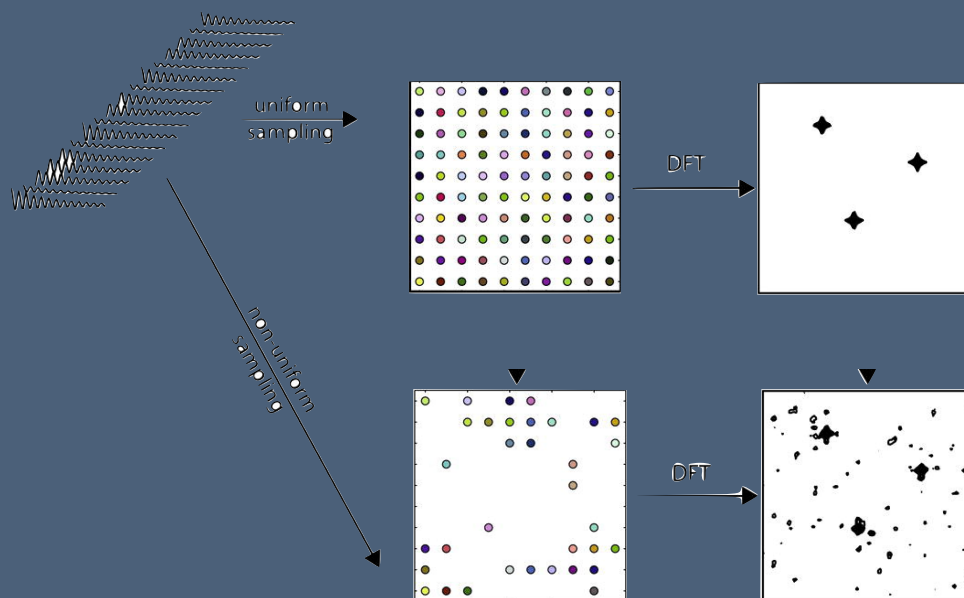
Over several decades ...

- Struggle to match human performance - intuitive human judgment
- Only on 'perfect datasets' - sharp well-resolved with minimal noise

Deep Learning

- Automate or improve different analysis stages
- Increasing efficiency, utility and ease of use of NMR spectroscopy

# Reconstruction of non-uniformly sampled (NUS)



Several excellent algorithms for reconstructing NUS NMR spectra using non-DL methods:

- SMILE (Ying et al. 2017)
- hmsIST (Hyberts et al. 2012)
- MDD-NMR (Jaravine et al. 2006).

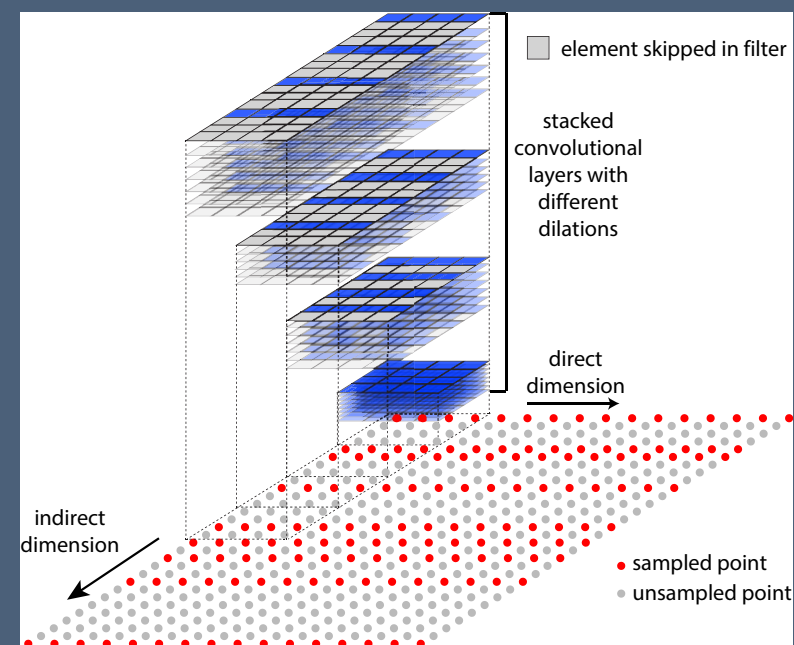
Recent proof-of-principles studies have shown that DL based reconstruction methods have the ability to give reconstructions more rapidly and with higher fidelity than existing methods

# FID-Net

A versatile DNN architecture

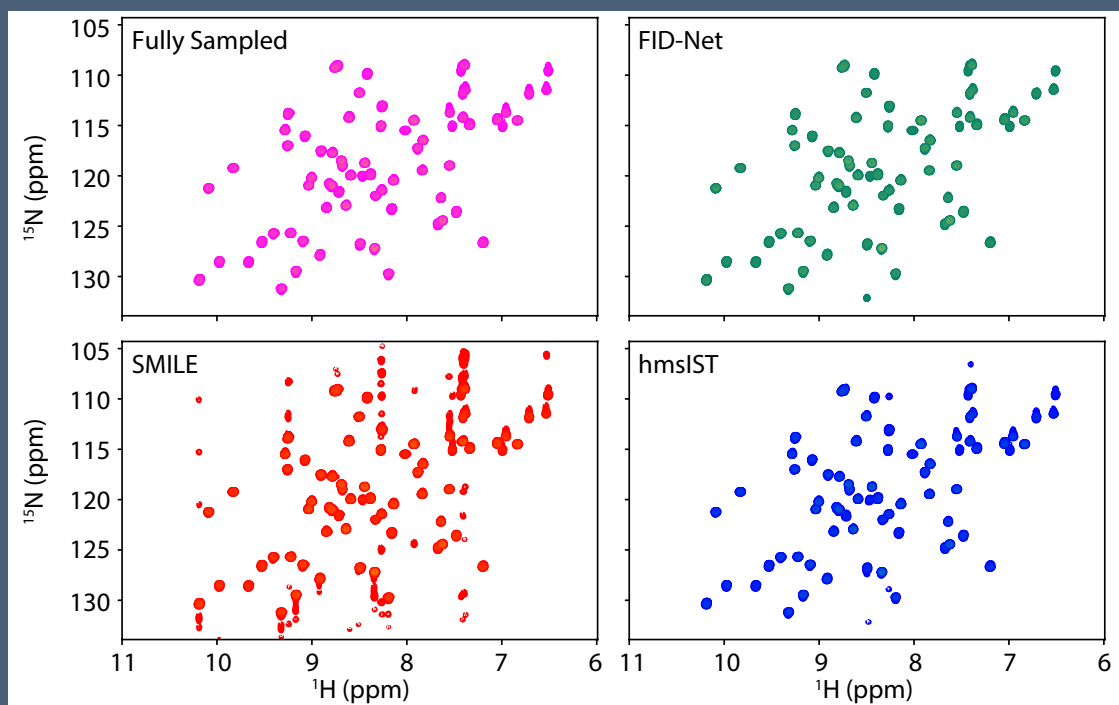
FID-Net works effectively:

- arbitrary sampling schedules (meaning can be deployed w/o further training & minimal user input)
- processing time domain data beyond reconstruction tasks
- Virtually decouple  $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$  couplings in HNCA spectra



# Reconstruction NUS spectra

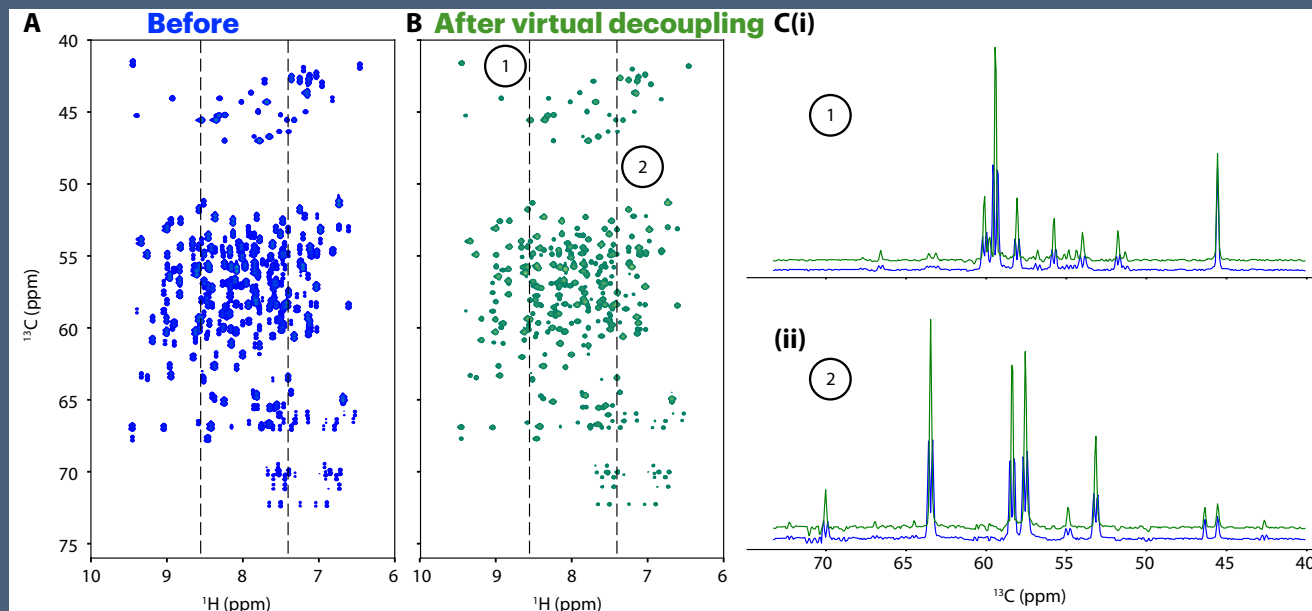
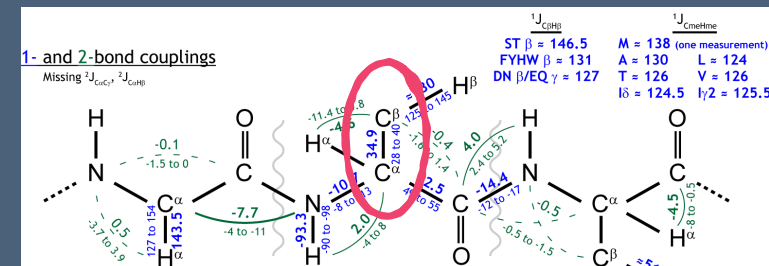
FID-Net vs. SMILE vs. hmsIST



- 12.5% sampling
- 100 sampling schedule with a different Poisson-gap sampling schedule

# Virtual Decoupling

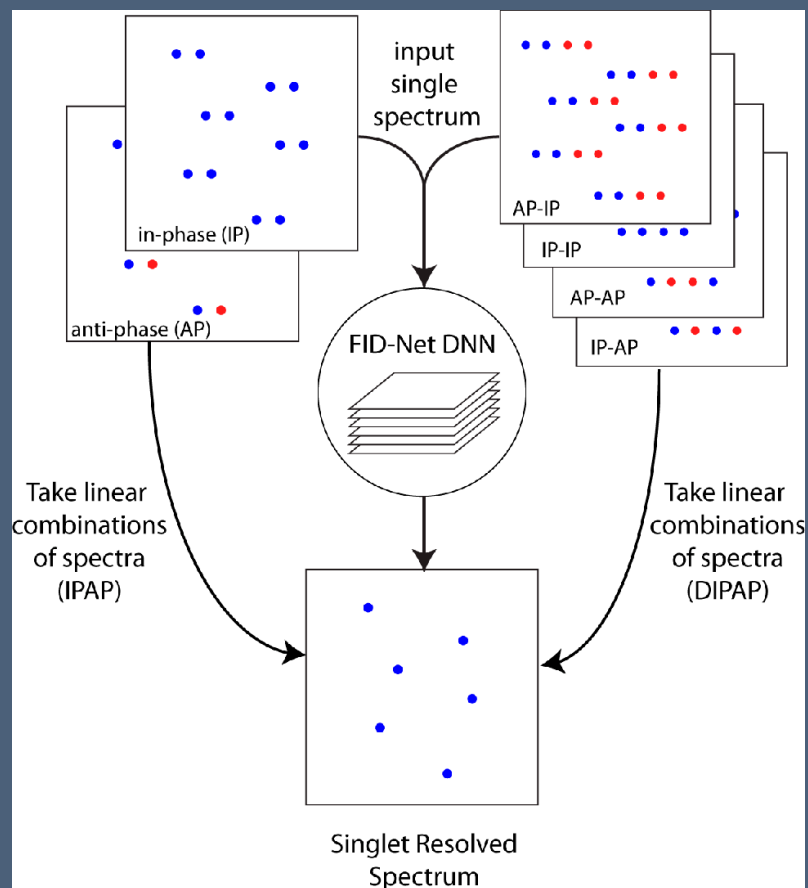
$^{13}\text{C}_\alpha$ - $^{13}\text{C}_\beta$  in 3D-HNCA and HN(CO)CA



How successful doublets are successfully decoupled yielding an improvement in resolution and two-fold increase in sensitivity for these peaks while the singlet glycine peaks are unaltered

# Virtual Decoupling (II)

$^{13}\text{C}$ - $^{13}\text{C}$ ,  $^{13}\text{CO}$ - $^{15}\text{N}$ ,  $^{13}\text{C}_\alpha$ - $^{13}\text{CO}$  ...



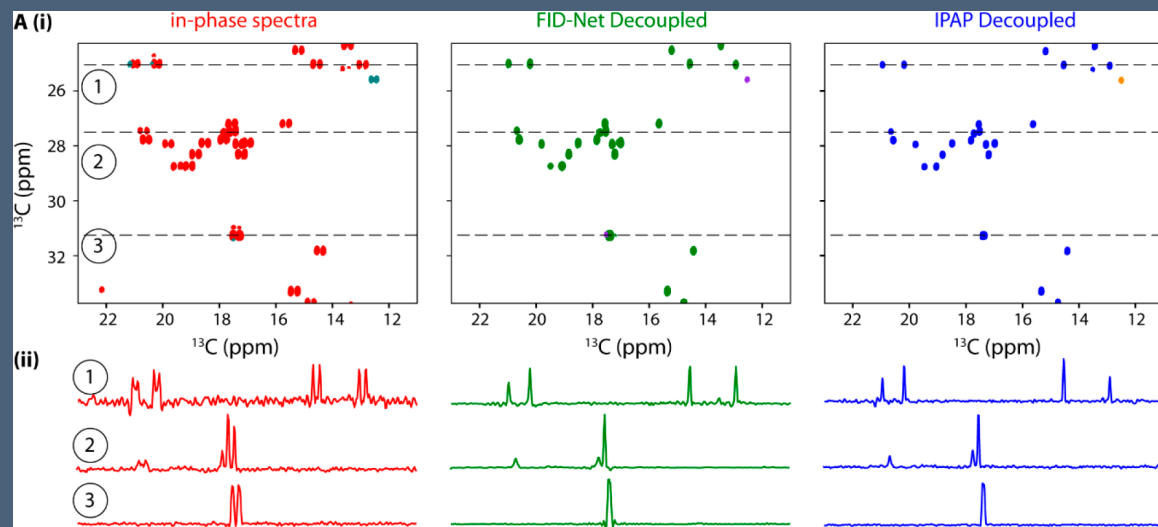
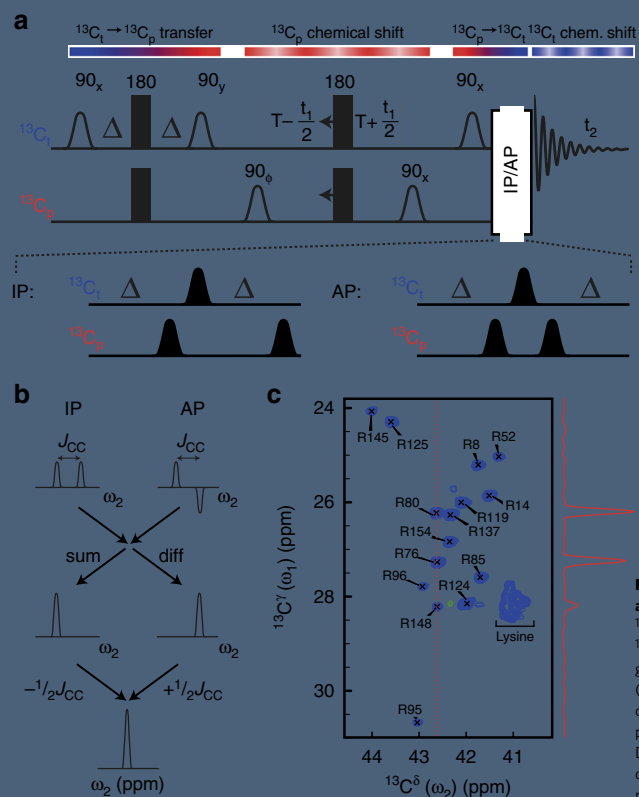
Traditional methods, such as IPAP and DIPAP, require the acquisition of multiple spectra and taking linear combinations to yield a singlet resolved spectrum.

Conversely, FID-Net-based DNNs can be trained to decouple spectra with one or two couplings using a single spectrum

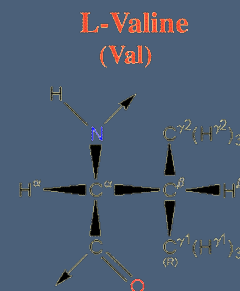


# Virtual Decoupling (III)

## $^{13}\text{C}$ - $^{13}\text{C}$ side-chain correlation spectra for per-deuterated proteins

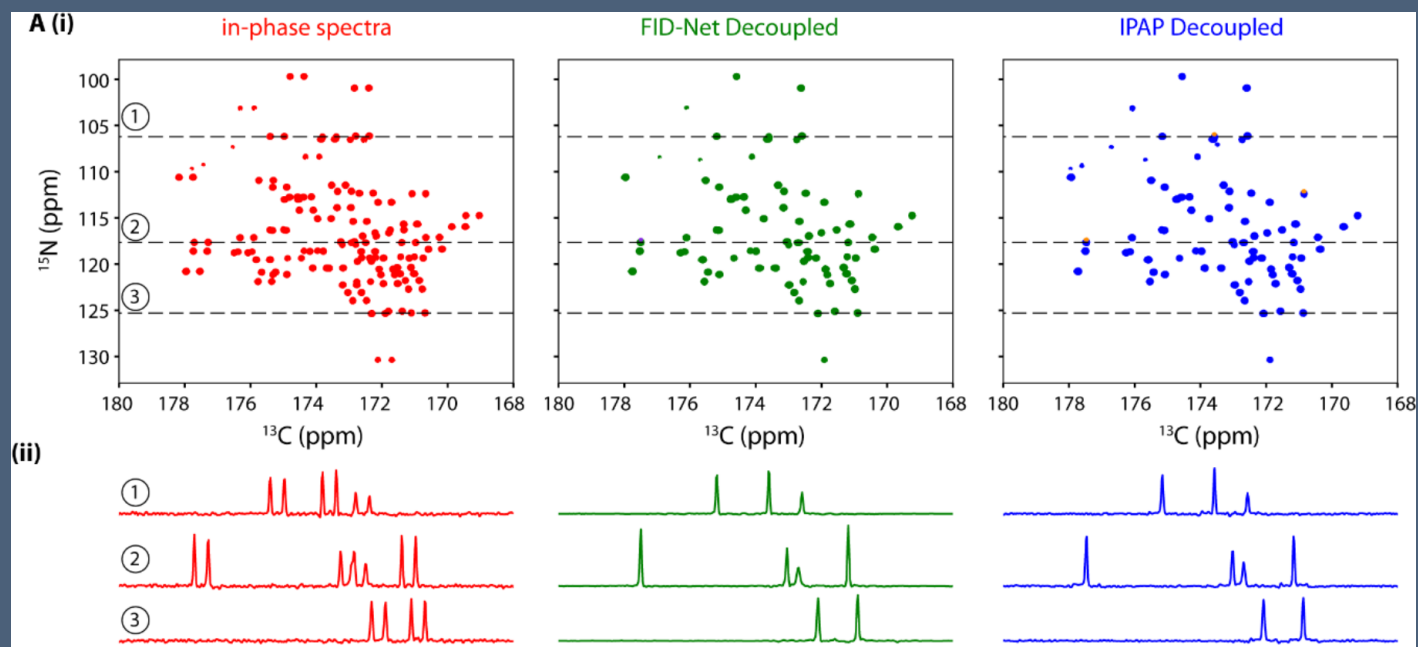


**Fig. 2**  $^{13}\text{C}$ - $^{13}\text{C}$  side-chain correlation spectra of per-deuterated proteins. **a** Schematic representation of the NMR pulse sequence used to obtain  $^{13}\text{C}$ - $^{13}\text{C}$  side-chain correlation spectra. The flow of the magnetisation between  $^{13}\text{C}_\alpha$  (blue) and  $^{13}\text{C}_\beta$  (red) is shown above the sequence with colour gradients. The following delays are used:  $\Delta = 1/(4J_{CC}) \approx 7.1$  ms,  $T = 1/(2J_{CC}) \approx 14.1$  ms, where  $J_{CC}$  is the one-bond  $^{13}\text{C}$ - $^{13}\text{C}$  scalar coupling constant. Rectangular pulses are high-power and not selective, bell-shaped pulses are frequency selective ( $90^\circ$ : white outlined,  $180^\circ$ : black). Deuterium,  $^2\text{H}$ , is decoupled throughout the sequence and frequency discrimination is obtained by states-TPPI of phase  $\phi^1$ . **b** Schematic representation of post-processing to obtain the decoupled spectrum. **c** Arginine  $^{13}\text{C}_\beta$ - $^{13}\text{C}_\gamma$  correlation of the 18-kDa protein T4L L99A, obtained on a 1.4 mM sample at a static field of 14.1 T at 278 K in 37 min



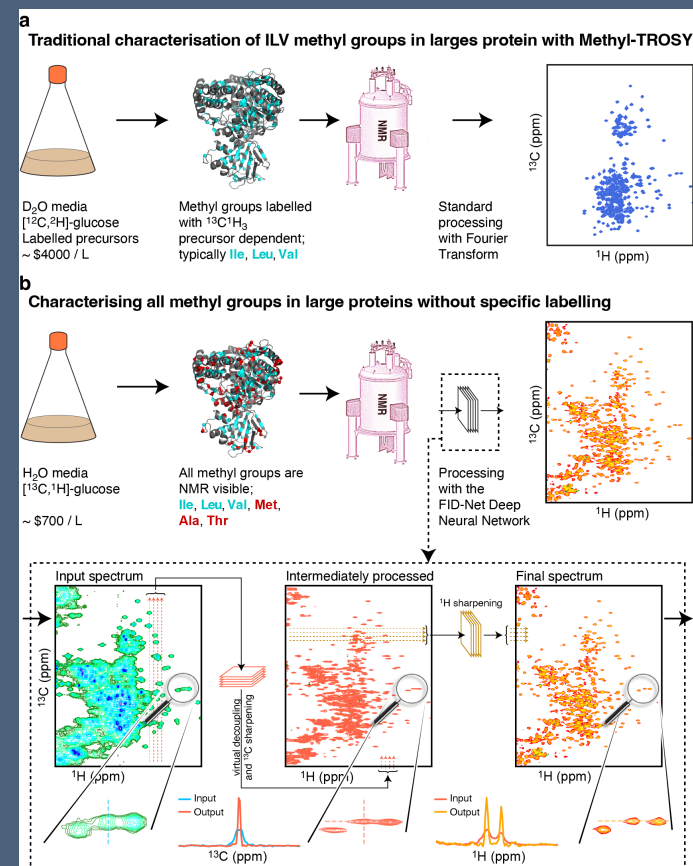
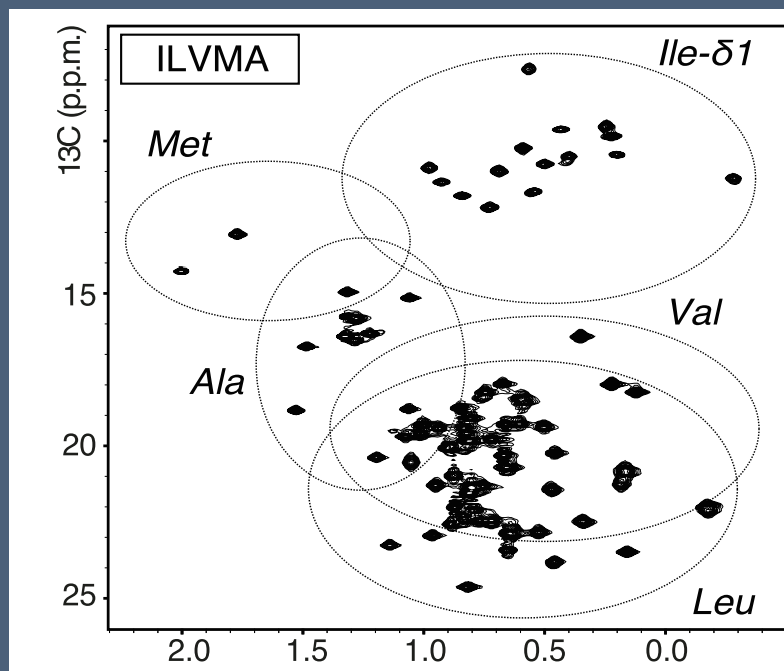
# Virtual Decoupling (IV)

$^{13}\text{C}$ - $^{15}\text{N}$  correlation spectra (CON)



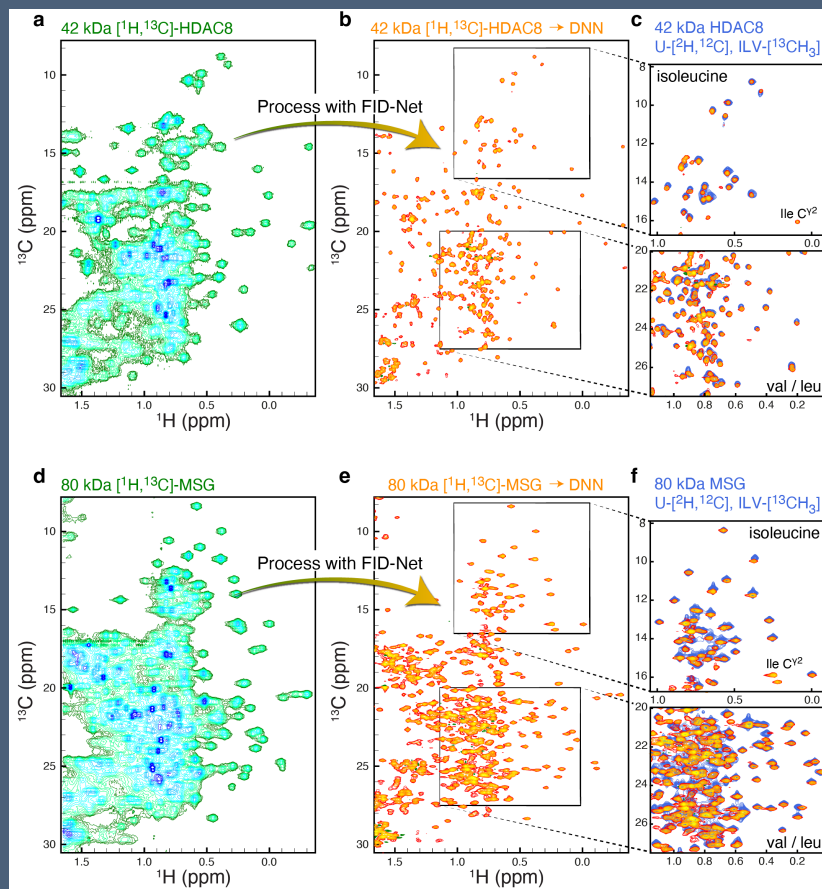
# FID-Net on large protein systems

## Methyl NMR Spectroscopy



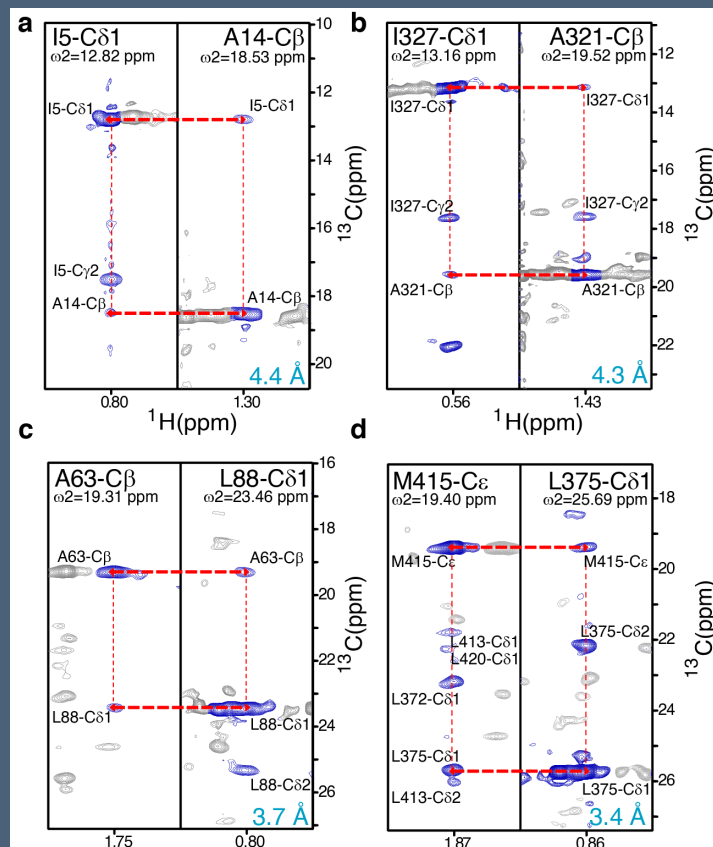
# FID-Net on large protein systems

## Methyl NMR Spectroscopy

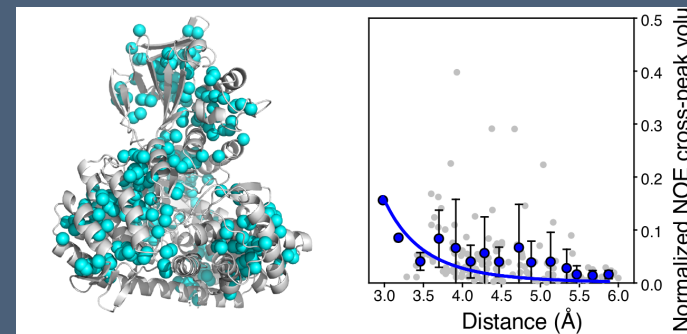


# FID-Net on large protein systems

3D  $^{13}\text{C}$ -HSQC-NOESY-HSQC non deuterated 80kDa MSG

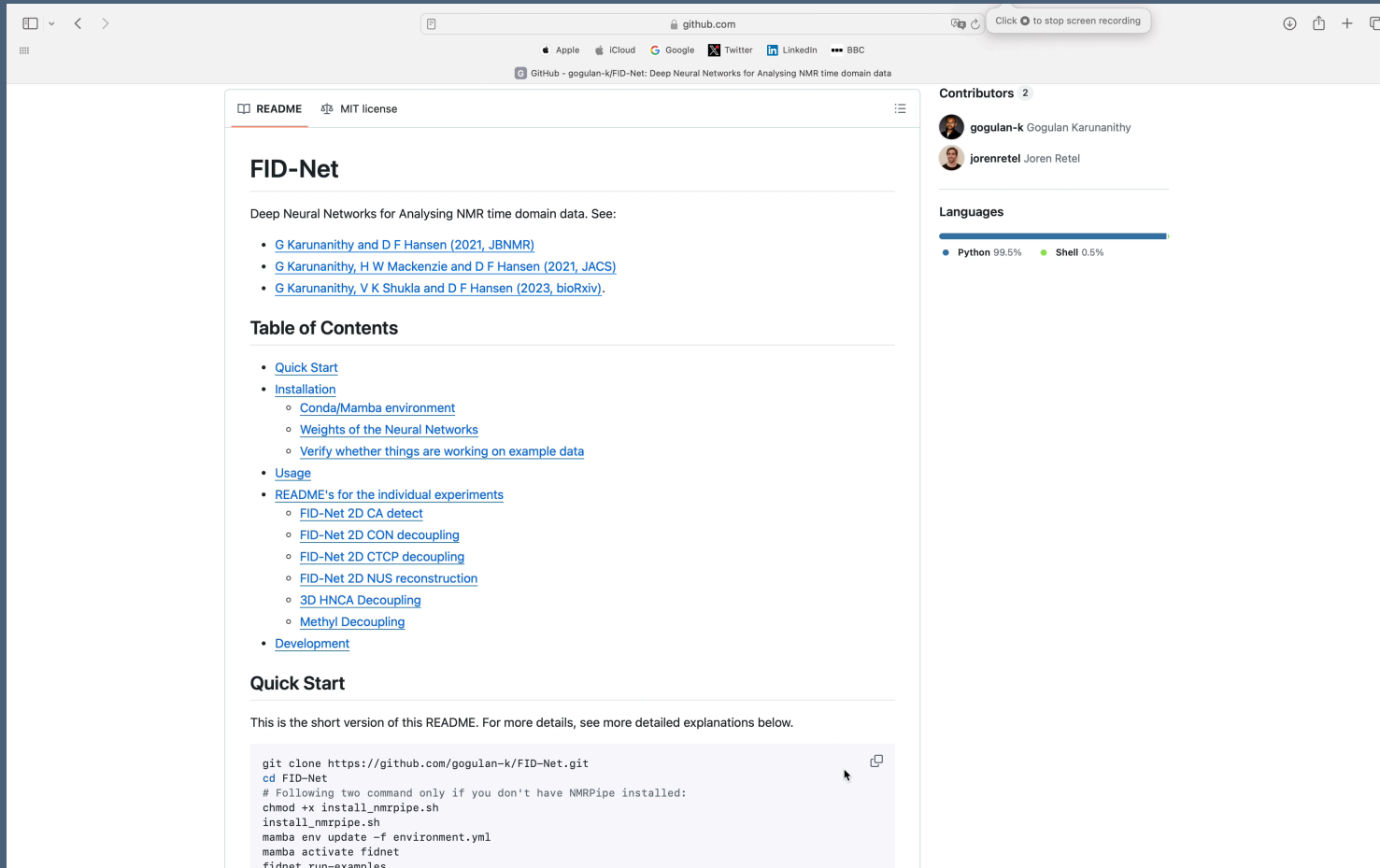


312 NOE cross-peaks were observed among 170 methyl bearing residues from different regions of the protein



# How to ...

- <https://github.com/gogulan-k/FID-Net>
- [d.hansen@ucl.ac.uk](mailto:d.hansen@ucl.ac.uk)



The screenshot shows the GitHub repository page for 'FID-Net'. The main content area displays the README, which includes a title 'FID-Net', a description 'Deep Neural Networks for Analysing NMR time domain data. See:', a list of references, a 'Table of Contents' with links to sections like 'Quick Start', 'Installation', 'Usage', and 'Development', and a 'Quick Start' section with a code block for cloning and running the project. The right sidebar shows 'Contributors' (gogulan-k and Jorenretel) and a 'Languages' bar indicating Python (99.5%) and Shell (0.5%).

**FID-Net**

Deep Neural Networks for Analysing NMR time domain data. See:

- [G Karunanithy and D F Hansen \(2021, JBNMR\)](#)
- [G Karunanithy, H W Mackenzie and D F Hansen \(2021, JACS\)](#)
- [G Karunanithy, V K Shukla and D F Hansen \(2023, bioRxiv\)](#)

**Table of Contents**

- [Quick Start](#)
- [Installation](#)
  - [Conda/Mamba environment](#)
  - [Weights of the Neural Networks](#)
  - [Verify whether things are working on example data](#)
- [Usage](#)
- [README's for the individual experiments](#)
  - [FID-Net 2D CA detect](#)
  - [FID-Net 2D CON decoupling](#)
  - [FID-Net 2D CTCP decoupling](#)
  - [FID-Net 2D NUS reconstruction](#)
  - [3D HNCA Decoupling](#)
  - [Methyl Decoupling](#)
- [Development](#)

**Quick Start**

This is the short version of this README. For more details, see more detailed explanations below.

```
git clone https://github.com/gogulan-k/FID-Net.git
cd FID-Net
# Following two command only if you don't have NMRPipe installed:
chmod +x install_nmrpipe.sh
install_nmrpipe.sh
mamba env update -f environment.yml
mamba activate fidnet
fidnet run-examples
```