



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

## FLORE

# Repository istituzionale dell'Università degli Studi di Firenze

### **13C APSY-NMR for sequential assignment of intrinsically disordered proteins**

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

*Original Citation:*

13C APSY-NMR for sequential assignment of intrinsically disordered proteins / Murralli, Maria Grazia; Schiavina, Marco; Sainati, Valerio; Bermel, Wolfgang; Pierattelli, Roberta; Felli, Isabella C. - In: JOURNAL OF BIOMOLECULAR NMR. - ISSN 0925-2738. - STAMPA. - 70:(2018), pp. 167-175. [10.1007/s10858-018-0167-4]

*Availability:*

The webpage <https://hdl.handle.net/2158/1119581> of the repository was last updated on 2021-03-21T10:37:03Z

*Published version:*

DOI: 10.1007/s10858-018-0167-4

*Terms of use:*

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (<https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf>)

*Publisher copyright claim:*

La data sopra indicata si riferisce all'ultimo aggiornamento della scheda del Repository FloRe - The above-mentioned date refers to the last update of the record in the Institutional Repository FloRe

(Article begins on next page)



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

**SBA**  
SISTEMA  
BIBLIOTECARIO  
DI ATENE0

*This is an author version of the contribution published on:*  
*Questa è la versione dell'autore dell'opera:*  
***<sup>13</sup>C apsy-nmr for sequential assignment of intrinsically disordered  
proteins  
in  
JOURNAL OF BIOMOLECULAR NMR, 70(3), 2018, p. 167-175]***

*The definitive version is available at:*  
*La versione definitiva è disponibile alla URL:*  
<https://doi.org/10.1007/s10858-018-0167-4>

**Coordinamento centrale biblioteche**

Via G. Capponi, 7 – 50121 Firenze

Tel. +39055.2756555-56, cell. +393385222146 · Fax +39055.2756332 ·

e-mail: [cb@sba.unifi.it](mailto:cb@sba.unifi.it) – posta certificata: [sba@pec.unifi.it](mailto:sba@pec.unifi.it)

P.IVA | Cod. Fis. 01279680480

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26

# **$^{13}\text{C}$ APSY-NMR for sequential assignment of intrinsically disordered proteins**

Maria Grazia Murrari<sup>1</sup>, Marco Schiavina<sup>1</sup>, Valerio Sainati<sup>1</sup>, Wolfgang Bermel<sup>2</sup>, Roberta Pierattelli<sup>1,3</sup>✉ and Isabella C. Felli<sup>1,3</sup>✉

<sup>1</sup> CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy

<sup>2</sup> Bruker BioSpin GmbH, Silberstreifen, 76287 Rheinstetten, Germany

<sup>3</sup> Department of Chemistry “Ugo Schiff”, University of Florence, 50019 Sesto Fiorentino, Italy

Corresponding Authors

E-mail: felli@cerm.unifi.it; pierattelli@cerm.unifi.it

1 **Abstract**

2 The increasingly recognized biological relevance of intrinsically disordered proteins requires  
3 a continuous expansion of the tools for their characterization via NMR spectroscopy, the  
4 only technique so far able to provide atomic-resolution information on these highly mobile  
5 macromolecules. Here we present the implementation of projection spectroscopy in  $^{13}\text{C}$ -  
6 direct detected NMR experiments to achieve the sequence specific assignment of IDPs. The  
7 approach was used to obtain the complete backbone assignment at high temperature of  $\alpha$ -  
8 synuclein, a paradigmatic intrinsically disordered protein.

9

10

11

12 **Keywords**

13 Intrinsically disordered proteins, IDPs, assignment, NUS,  $^{13}\text{C}$  detection

14

15

## 1 **Introduction**

2 The functional importance of intrinsically disordered proteins (IDPs) or protein regions  
3 (IDPRs) has now been widely recognized (Habchi et al., 2014; Wright & Dyson, 2015; van  
4 der Lee et al., 2014). Furthermore, a number of key proteins related to the onset of diseases  
5 is intrinsically disordered or has large disordered regions (Uversky et al., 2008). Perhaps the  
6 most well-known examples include proteins involved in the development of  
7 neurodegenerative diseases, such as  $\alpha$ -synuclein and Tau for Parkinson's and Alzheimer's  
8 diseases. In addition, a vast range of examples of IDPs/IDPRs related to cancer progression  
9 are emerging, beyond the most well-known example of p53. These include c-myc, p21, AR,  
10 BRCA to name just a few recent examples (Uversky et al., 2014). Therefore, novel  
11 approaches to target IDPs or IDPRs are urgently needed in the field of drug discovery (Tóth  
12 et al., 2014; Ambadipudi & Zweckstetter, 2016; Joshi et al., 2016; Heller et al., 2017) and in  
13 this frame it becomes very important to develop NMR methods to study IDPs/IDPRs in their  
14 native state under physiological conditions (Felli and Pierattelli eds., 2015).

15 The peculiar amino acidic composition of IDPs, often characterized by multiple sequence  
16 repetitions and low complexity regions, the high flexibility typical of IDPs/IDPRs and the  
17 many conformations sampled in which the protein backbone is largely solvent exposed have  
18 an impact on NMR observables that should be considered in the design of effective  
19 experimental NMR approaches (Brutscher et al., 2015).

20 The critical points that need to be faced in order to optimize NMR spectra for the study of  
21 highly flexible IDPs/IDPRs include the very low chemical shift dispersion typical of proteins  
22 that lack a stable, well defined 3D structure, as well as solvent exchange processes that  
23 broaden amide proton resonances when approaching physiological conditions. As known  
24 since the early studies on protein folding, heteronuclear spins are characterized by a large  
25 chemical shift dispersion and thus well suited to characterize disordered proteins (Dyson &  
26 Wright, 2001). Indeed, initial studies of unfolded proteins were stimulated by the  
27 development of 3D triple resonance experiments (Kay et al., 1990; Sattler et al., 1999).  
28 However, to study unfolded/intrinsically disordered proteins of increasing complexity, a  
29 further push in resolution was necessary; a major contribution towards this goal derived from  
30 the extension of the dimensionality of NMR experiments to more than three dimensions. Of  
31 course methods based on uniform sampling of indirect dimensions and on conventional  
32 processing strategies were not suitable to achieve a high resolution in all dimensions, a key  
33 aspect to study complex IDPs/IDPRs; different approaches needed to be developed to

1 render these experiments feasible in a reasonable amount of time and manageable in terms  
2 of processing time and disk space. Many non-uniform sampling approaches and different  
3 strategies to process the data were indeed proposed (Kupce & Freeman, 2004; Kim &  
4 Szyperski, 2003; Hiller et al., 2005; Hoch & Stern, 2001; Orekhov et al., 2003; Kazimierczuk  
5 et al., 2006) and initially implemented in triple resonance multidimensional experiments  
6 based on  $H^N$  detection (Narayanan et al., 2010; Fiorito et al., 2006; Zawadzka-Kazimierczuk  
7 et al., 2010; Motackova et al., 2010) and on  $H^\alpha$  detection (Yao et al., 2014, Mäntylähti et al.,  
8 2011) showing a great potential for the characterization of IDPs (Fiorito et al., 2006;  
9 Zawadzka-Kazimierczuk et al., 2012; Narayanan et al., 2010; Yao et al., 2014).

10 In parallel, the increase of the sensitivity of NMR instrumentation (Kovacs et al., 2005)  
11 enabled the development of experimental strategies based on heteronuclear direct detection  
12 (Bermel et al., 2006c), which turned out to be an excellent tool to study IDPs. Indeed the  
13 valuable chemical shift dispersion of heteronuclear spins can be exploited in all dimensions  
14 of NMR experiments, including the directly detected one, and limitations deriving from  
15 solvent exchange broadening are obviously reduced. Several methods to overcome the  
16 problem of  $^{13}C$  homonuclear couplings in the direct acquisition dimension were proposed  
17 (Felli & Pierattelli, 2015) and enabled the application of  $C'$  direct detection with high  
18 resolution. A suite of 3D multidimensional  $C'$  detected experiments was proposed (Bermel  
19 et al., 2009a; Bermel et al., 2009b; Felli et al., 2009; Bermel et al., 2008; Bermel et al.,  
20 2006b, O'Hare et al., 2009) and proved to be very useful for sequence specific assignment  
21 and characterization of several IDPs. The suite was then expanded to include also  $C'$   
22 detected experiments of dimensionality higher than 3 by implementing random non-uniform  
23 sampling in combination with sparse multidimensional Fourier transform (SMFT) processing  
24 of the data showing that experiments of higher dimensionality based on  $C'$  detection are  
25 indeed possible and constitute a valuable tool to study IDPs (Nováček et al., 2011; Bermel  
26 et al., 2012; Novacek et al., 2012; Haba et al., 2013; Bermel et al., 2013; Piai et al., 2014;  
27 Baias et al., 2017).

28 As a further step facilitating the exploitation of the  $^{13}C$ -detection based assignment strategy,  
29 we present the implementation of projection spectroscopy for  $C'$  detected multidimensional  
30 NMR experiments. The approach is tested on  $\alpha$ -synuclein at high temperature.

31

32

## 1 **Material and Methods**

2 Isotopically enriched  $\alpha$ -synuclein was prepared as previously described (Huang, Ren et al.,  
3 2005) A sample of 1.0 mM uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labeled human  $\alpha$ -synuclein in 200 mM NaCl,  
4 0.5 mM EDTA, 20 mM phosphate buffer at pH 6.5 was used. 10 %  $\text{D}_2\text{O}$  was added for the  
5 lock.

6 All the 2D NMR experiments were acquired at 285.5 K, 295 K, 305 K and 315 K while the  
7 APSY experiments were all acquired at 315 K on a Bruker AVANCE NEO spectrometer  
8 operating at 700.06 MHz  $^1\text{H}$ , 176.05 MHz  $^{13}\text{C}$  and 70.97 MHz  $^{15}\text{N}$  frequencies equipped with  
9 a cryogenically cooled probehead optimized for  $^{13}\text{C}$ -direct detection (TXO). Carrier  
10 frequencies and RF pulses suitable for triple resonance  $^{13}\text{C}$  detected experiments were used  
11 and are summarized hereafter while specific parameters for the different experiments are  
12 reported in Table 1. Q5 and Q3 shapes (Emsley & Bodenhausen, 1992) of durations of 300  
13 and 231  $\mu\text{s}$ , respectively, were used for  $^{13}\text{C}$  band-selective  $\pi/2$  and  $\pi$  flip angle pulses,  
14 except for the  $\pi$  pulses that should be band-selective on the  $\text{C}^\alpha$  region (Q3, 1200  $\mu\text{s}$ ) and  
15 for the adiabatic  $\pi$  pulse to invert both  $\text{C}'$  and  $\text{C}^\alpha$  (smoothed Chirp 500  $\mu\text{s}$ , 20 % smoothing,  
16 80 kHz sweep width, 11.3 kHz RF field strength) (Böhlen & Bodenhausen, 1993). The  $^{13}\text{C}$   
17 band selective pulses on  $\text{C}^\alpha$  and  $\text{C}'$  were applied at the center of each region, 53 and 173.5  
18 ppm respectively; the Chirp pulse was centered at 90 ppm. Carrier frequencies for  $^{15}\text{N}$  and  
19  $^1\text{H}$  were 122.5 and 4.7 ppm respectively. Decoupling of  $^1\text{H}$  and  $^{15}\text{N}$  was achieved with Waltz  
20 (2.6 kHz) and Garp-4 (1.0 kHz) decoupling sequences respectively (Shaka et al., 1985). All  
21 gradients employed had a smoothed square shape. The parameters used for the acquisition  
22 of the 5D and 4D experiments as well as the parameters selected to implement the APSY  
23 approach are reported in Tables 1 and 2 respectively. FLOPSY-16 (Kadkhodaie et al, 1991)  
24 was used in the 4D experiment to spin-lock. All the spectra were acquired using Bruker  
25 TopSpin 4.0.1 software. Calibration of the spectra was achieved using DSS as a standard  
26 for  $^1\text{H}$  and  $^{13}\text{C}$ ;  $^{15}\text{N}$  shifts were calibrated indirectly (Markley et al., 1998).

27

28

## 1 **Results and discussion**

2 The 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC and 2D  $^{13}\text{C}$ - $^{15}\text{N}$  CON experiments of  $\alpha$ -synuclein recorded at various  
3 temperatures are reported in Figure 1. It can be noted that the number of cross peaks that  
4 can be observed through 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra decreases with increasing temperature.  
5 This is the result of the pronounced exchange processes of amide protons with water  
6 protons due to the largely solvent exposed backbone of  $\alpha$ -synuclein.

7 Nuclear spins of non-exchangeable atoms should thus be exploited in NMR experiments.  
8 However, even if in principle very sensitive,  $^1\text{H}$  nuclear spins of aliphatic and aromatic  
9 residues are characterized by a quite narrow chemical shift dispersion, in particular when  
10 considering amino acids of the same (or similar) type in proteins lacking a 3D structure such  
11 as in IDPs. In addition the  $^1\text{H}$ - $^1\text{H}$  homonuclear couplings, which are not easy to remove in  
12 the direct acquisition dimension, constitute an additional drawback when really high  
13 resolution is needed. Instead, non-exchangeable heteronuclear spins, such as  $^{13}\text{C}$  and  $^{15}\text{N}$ ,  
14 provide a large chemical shift dispersion even if at the expense of reduced sensitivity when  
15 directly detected. Several strategies can be implemented for  $\text{C}'$  homonuclear decoupling  
16 (Felli and Pierattelli, 2015). Thus, correlations between heteronuclear spins involved in the  
17 peptide bond are the most appropriate ones to achieve high resolution (Bermel et al., 2013),  
18 as evidenced in the spectra reported in Figure 1 which could be acquired with just a few  
19 scans per increment; 2D  $^{13}\text{C}$ - $^{15}\text{N}$  CON spectra (Bermel et al., 2006a) thus provide a valuable  
20 tool to study IDPs at physiological conditions (Gil et al., 2013).

21 The 2D  $^{13}\text{C}$ - $^{15}\text{N}$  CON spectra can then be expanded into higher dimensional experiments  
22 that provide different types of correlations between different nuclear spins in order to achieve  
23 sufficient information for sequence specific assignment. Increasing the number of  
24 dimensions however, even when exploiting heteronuclear spins to take advantage of their  
25 high dispersion in IDPs, is effective only if we are able to achieve a very high resolution in  
26 the indirectly sampled dimensions. This is practically not feasible due to time constraints as  
27 well as disk usage unless non-uniform sampling strategies are used, which, combined with  
28 appropriate processing strategies, have shown to be very effective (Bermel et al., 2013;  
29 Bermel et al., 2012; Novacek et al., 2012; Novacek et al., 2011). Here we test the  
30 applicability of projection spectroscopy and of the relative automated data analysis (Hiller et  
31 al., 2005) to the  $\text{C}'$  detected 5D (HCA)CONCACON experiment (Bermel et al., 2013), which  
32 correlates the CON peak of one peptide bond with the preceding one, for the assignment of  
33  $\alpha$ -synuclein at high temperature (315 K).

1 The four orthogonal planes that can be collected through this experiment, reported in Figure  
2 2, show the excellent resolution that can be achieved in all dimensions thanks to the  
3 contribution of heteronuclear chemical shifts, in particular of C' and N that provide  
4 correlations across one peptide bond (A) and of C' and N belonging to the same amino acid  
5 (B), as well as sequential C'-C' correlations (C). The plane correlating C' and C<sup>α</sup> (D) is  
6 characterized by a somehow reduced resolution but still provides useful information about  
7 the amino acid type. These are excellent requisites to implement projection spectroscopy  
8 through which not only orthogonal planes but also “transverse” planes cutting the  
9 multidimensional object at different angles are acquired with very high resolution as shown  
10 in panels E and F of Figure 2. Thanks to the excellent resolution that can be obtained in the  
11 projections of C' detected exclusively heteronuclear experiments, it is possible to resolve  
12 the majority of cross peaks in the projections and to identify essentially all correlations  
13 expected in the 5D spectrum just by inspection of a limited set of projections. In the present  
14 case, 64 projections were acquired with very high resolution (256 points were set in each of  
15 the 4 indirect dimensions); it is worth noting that to achieve this kind of resolution for the  
16 analogous 5D experiment would have required a prohibitively long time (hundreds of years!).  
17 The analysis of the 64 projections, which can be performed automatically by using the  
18 algorithm GAPRO (Geometric Analysis of PROjections) (Hiller, Wider et al., 2008; Hiller &  
19 Wider, 2012) as it is implemented in the TopSpin 4.0.1 software, provides 191 5D  
20 correlations (131 C'<sub>i-1</sub>, N<sub>i</sub>, C<sup>α</sup><sub>i</sub>, C'<sub>i</sub>, N<sub>i+1</sub> and 60 C'<sub>i-1</sub>, N<sub>i</sub>, C<sup>α</sup><sub>i-1</sub>, C'<sub>i-1</sub>, N<sub>i</sub>). These correlations  
21 constitute an excellent tool to achieve sequence specific assignment by linking the shifts of  
22 C' and N involved in a peptide bond with those of the previous one, and also giving some  
23 hint on the kind of amino acid through the information of the C<sup>α</sup> chemical shift. The  
24 correlations identified through the GAPRO analysis of the projections are stored in the form  
25 of N-dimensional peak-lists that can be manually inspected by loading them on the different  
26 projections. In particular, by loading the peak-lists on a 2D <sup>13</sup>C-<sup>15</sup>N CON spectrum the  
27 completeness of the assignment can be verified and it is possible to perform the sequence  
28 specific walk through the backbone just by moving from one cross peak (C'<sub>i</sub>, N<sub>i+1</sub>) to the  
29 neighboring one (C'<sub>i-1</sub>, N<sub>i</sub>), as shown schematically in Figure 3 for a few residues. It is worth  
30 noting that for most of these residues the H<sup>N</sup> resonance was not observable in <sup>1</sup>H detected  
31 experiments.

32 A few comments are due on technical details related to the implementation of the APSY  
33 approach for the 5D (HCA)CONCACON experiment (Bermel et al., 2013). In contrast to the  
34 original sequence, the t<sub>2</sub> period was implemented in a semi-constant time manner in order

1 to increase to possible resolution in this dimension. As noted above, the version presently  
2 used yields two types of correlations: type I ( $C'_{i-1}$ ,  $N_i$ ,  $C^\alpha_i$ ,  $C'_i$ ,  $N_{i+1}$ ) which contains the  
3 information for sequential assignment and type II ( $C'_{i-1}$ ,  $N_i$ ,  $C^\alpha_{i-1}$ ,  $C'_{i-1}$ ,  $N_i$ ) which correlate  
4 three nuclear spins ( $C^\alpha_{i-1}$ ,  $C'_{i-1}$ ,  $N_i$ ) giving essentially the same information that can be found  
5 in 3D CACON experiments. The experiment is optimized to detect correlations of type I;  
6 correlations of type II can in principle be suppressed as suggested for the 3D  
7 (HCA)CON(CA)H (Mäntylähti et al. 2011). However, in the present case this subset of peaks  
8 could be easily identified directly from the peak list and further used for confirming the  
9 assignment, so that no further modification of the experiment was required. A few type I  
10 correlations were instead missing. Manual inspection of the peak list revealed that most of  
11 them involve residues in the known  $\alpha$ -synuclein repeats which could not all be safely  
12 identified through the automatic algorithm.

13 A second comment concerns the number of projections that need to be acquired to obtain  
14 the information for sequence specific assignment. In the present case we collected  
15 projections at 30 different combinations of angles  $\alpha$ ,  $\beta$ ,  $\gamma$  that, considering the issue of  
16 quadrature detection in all indirect dimensions, translates into 64 different 2D projections  
17 (Table 2). However, processing of the data using a smaller number of projection angles  
18 shows that the information content is essentially the same if we use only 26 combinations  
19 of angles, as suggested by the software (Hiller & Wider, 2012; Hiller, Wider et al., 2008),  
20 and that the number of type I correlations detected in the spectra is only slightly reduced if  
21 we consider a smaller subset of angles (Table 3), showing the robustness of this approach.  
22 The major issue to be solved was to properly implement in the APSY software the virtual  
23 homonuclear decoupling of  $C'$  through the IPAP approach. The processing of the data,  
24 including the virtual decoupling, can now be handled very easily through TopSpin 4.0.1.

25 The APSY approach can of course be implemented in any kind of  $C'$  detected experiment,  
26 using the 2D  $^{13}\text{C}$ - $^{15}\text{N}$  CON spectrum as a reference. For example, the information about the  
27 amino acid side chain, can be easily obtained and correlated to each of the cross peaks  
28 observed in the 2D  $^{13}\text{C}$ - $^{15}\text{N}$  CON spectrum, by collecting experiments such as 4D  
29 HCBCACON (Bermel et al., 2009) or 4D HCCCON (Bermel et al., 2012). As an example the  
30 latter was acquired on  $\alpha$ -synuclein and provided information about the vast majority of  $^{13}\text{C}$   
31 chemical shifts of aliphatic side chains. Care was taken to implement the  $^{13}\text{C}$  evolution in  
32 the constant-time mode in order to refocus homonuclear  $^{13}\text{C}$  couplings of aliphatic carbons  
33 that easily show up in highly resolved projections and can complicate the analysis of

1 crowded regions. The modified pulse sequence (4D CT-HCCCON) and a few projections to  
2 illustrate the quality of the spectra are reported in the Supplementary Material (Figures S1  
3 and S2).

4 A final comment is due about the amide proton resonances that are still observable  
5 approaching physiological conditions. Combination of the C' detected experiments with H<sup>N</sup>  
6 detected ones provides a useful tool to achieve unambiguous assignment of the amide  
7 proton resonances that are still observable in these conditions. In our case we collected a  
8 3D HNCO experiment, also in the APSY mode, in order to correlate H<sup>N</sup><sub>i</sub> resonances with the  
9 C'<sub>i-1</sub>-N<sub>i</sub> resonances that we have assigned through C' detected experiments. Through this  
10 approach 93 cross peaks were detected and, thanks to the peak list obtained in a fully  
11 automated manner through the GAPRO analysis of the APSY HNCO, the available  
12 assignment of heteronuclear spins could easily be extended to the residual H<sup>N</sup> resonances  
13 that can be detected at this temperature. It is worth noting that a similar result could be  
14 achieved by following the temperature dependence of the signals. The latter approach  
15 however works well only for well resolved H<sup>N</sup> cross peaks but it becomes quite ambiguous  
16 for cross peaks in crowded regions of the 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectra. Therefore the  
17 combination of C' detected experiments with H<sup>N</sup> detected ones, even at high temperature  
18 and pH, provides unambiguous information to assign H<sup>N</sup> resonances that remain also in  
19 presence of fast exchange, without ambiguities resulting from temperature dependence of  
20 the signals and/or from incomplete information provided by H<sup>N</sup> detected spectra. The  
21 temperature dependence of H<sup>N</sup>, C' and N spins, once unambiguous assignment is available,  
22 might provide also valuable information to investigate variations of exchange processes of  
23 amide protons with the solvent upon increasing temperature as well as to highlight possible  
24 variations in partially populated local conformations by analyzing heteronuclear chemical  
25 shifts. The sequence specific assignment of α-synuclein achieved at 315 K has been  
26 deposited in the BMRB (ID 27348).

27 Concluding, we demonstrate the implementation of the APSY strategy in C' detected  
28 multidimensional NMR experiments. Excellent resolution can be achieved for the 2D cross  
29 sections of the high dimensional C' detected exclusively heteronuclear spectra, showing  
30 that the two approaches (APSY and C' detection) mutually benefit by their joint  
31 implementation.

32

33

1 **Acknowledgments**

2 The support and the use of resources of the CERM/CIRMMP center of Instruct-ERIC, a  
3 Landmark ESFRI project, is gratefully acknowledged. This work has been supported in part  
4 by a grant of the Fondazione CR di Firenze and by MEDINTECH (CTN01\_001177\_962865).

5

6

7

## 1   **References**

2

3   Ambadipudi S, Zweckstetter M (2016) Targeting intrinsically disordered proteins in rational drug discovery.  
4   Expert Opin Drug Discov 11:65-77

5   Baias M, Smith PE, Shen K, Joachimiak LA, Zerko S, Koźmiński W, Frydman J, Frydman L (2017) Structure  
6   and dynamics of the Huntingtin exon-1 N-terminus: a solution NMR perspective. J Am Chem Soc 139:1168-  
7   1176

8   Bermel W, Bertini I, Csizmok V, Felli IC, Pierattelli R, Tompa P (2009) H-start for exclusively heteronuclear  
9   NMR spectroscopy: the case of intrinsically disordered proteins. J Magn Reson 198:275-281

10   Bermel W, Bertini I, Felli IC, Kümmerle R, Pierattelli R (2006) Novel <sup>13</sup>C direct detection experiments, including  
11   extension to the third dimension, to perform the complete assignment of proteins. J Magn Reson 178:56-64

12   Bermel W, Bertini I, Felli IC, Lee Y-M, Luchinat C, Pierattelli R (2006) Protonless NMR experiments for  
13   sequence-specific assignment of backbone nuclei in unfolded proteins. J Am Chem Soc 128:3918-3919

14   Bermel W, Bertini I, Felli IC, Piccioli M, Pierattelli R (2006) <sup>13</sup>C-detected protonless NMR spectroscopy of  
15   proteins in solution. Progr NMR Spectrosc 48:25-45

16   Bermel W, Bertini I, Felli IC, Pierattelli R (2009) Speeding up <sup>13</sup>C direct detection Biomolecular NMR  
17   experiments. J Am Chem Soc 131:15339-15345

18   Bermel W, Bertini I, Gonnelli L, Felli IC, Koźmiński W, Piai A, Pierattelli R, Stanek J (2012) Speeding up  
19   sequence specific assignment of IDPs. J Biomol NMR 53:293-301

20   Bermel W, Bruix M, Felli IC, Kumar VMV, Pierattelli R, Serrano S (2013) Improving the chemical shift dispersion  
21   of multidimensional NMR spectra of intrinsically disordered proteins. J Biomol NMR 55:231-237

22   Bermel W, Felli IC, Gonnelli L, Koźmiński W, Piai A, Pierattelli R, Zawadzka-Kazimierczuk A (2013) High-  
23   dimensionality <sup>13</sup>C direct-detected NMR experiments for the automatic assignment of intrinsically disordered  
24   proteins. J Biomol NMR 57:353-361

25   Bermel W, Felli IC, Kümmerle R, Pierattelli R (2008) <sup>13</sup>C direct-detection biomolecular NMR. Concepts Magn  
26   Reson 32A:183-200

27   Böhlen J-M, Bodenhausen G (1993) Experimental aspects of chirp NMR spectroscopy. J Magn Reson Ser A  
28   102:293-301

29   Brutscher B, Felli IC, Gil-Caballero S, Hošek T, Kümmerle R, Piai A, Pierattelli R, Sólyom Z (2015) NMR  
30   Methods for the study of intrinsically disordered proteins structure, dynamics, and interactions: General  
31   overview and practical guidelines. Adv Exp Med Biol 870:122

32   Dyson HJ, Wright PE (2001) Nuclear magnetic resonance methods for the elucidation of structure and  
33   dynamics in disordered states. Methods In Enzymology 339:258-271

34   Emsley L, Bodenhausen G (1992) Optimization of shaped selective pulses for NMR using a quaternion  
35   description of their overall propagators. J Magn Reson 97:135-148

36   Felli IC, Pierattelli R (Eds), Intrinsically disordered proteins studied by NMR spectroscopy. Springer Switzerland

37   Felli IC, Pierattelli R (2015) Spin-state-selective methods in solution- and solid-state biomolecular <sup>13</sup>C NMR.  
38   Prog NMR Spectrosc 84:1-13

39   Felli IC, Pierattelli R, Glaser SJ, Luy B (2009) Relaxation-optimised Hartmann-Hahn transfer for carbonyl-  
40   carbonyl correlation spectroscopy using a specifically tailored MOCCA-XY16 mixing sequence for protonless  
41   <sup>13</sup>C direct detection experiments. J Biomol NMR 43:187-196

- 1 Fiorito F, Hiller S, Wider G, Wüthrich K (2006) Automated resonance assignment of proteins: 6D APSY-NMR.  
2 J Biomol NMR 35:27-37
- 3 Gil S, Hošek T, Solyom Z, Kümmerle R, Brutscher B, Pierattelli R, Felli IC (2013) NMR studies of intrinsically  
4 disordered proteins near physiological conditions. Angew Chem Int Ed 52:11808-11812
- 5 Haba NY, Gross R, Nováček J, Shaked H, Židek L, Barda-Saad M, Chill JH (2013) NMR determines transient  
6 structure and dynamics in the disordered C-terminal domain of WASp interacting protein. Biophysical J  
7 105:481-493
- 8 Habchi J, Tompa P, Longhi S, Uversky VN (2014) Introducing protein intrinsic disorder. Chem Rev 114:6561-  
9 6588
- 10 Hiller S, Fiorito F, Wüthrich K, Wider G (2005) Automated projection spectroscopy (APSY). Proc Natl Acad Sci  
11 USA 102:10876-10881
- 12 Hiller S, Wider G (2012) Automated projection spectroscopy and its applications. Top Curr Chem 316:21-47
- 13 Hiller S, Wider G, Wüthrich K (2008) APSY-NMR with proteins: practical aspects and backbone assignment.  
14 J Biomol NMR 42:179-195
- 15 Hoch JC Stern AS (2001) Nuclear Magnetic Resonance of Biological Macromolecules. pp 159-178
- 16 Huang C, Ren G, Zhou H, Wang C (2005) A new method for purification of recombinant human alpha-synuclein  
17 in Escherichia coli. Protein Expr Purif 42:173-177
- 18 Joshi P, Chia S, Habchi J, Knowles TPJ, Dobson CM, Vendruscolo M (2016) A fragment-based method of  
19 creating small-molecule libraries to target the aggregation of intrinsically disordered proteins. ACS Comb Sci  
20 18:144-153
- 21 Kadkhodaie M, Rivas O, Tan M, Mohebbi A, Shaka AJ, (1991) Broadband homonuclear cross polarization  
22 using flip-flop spectroscopy, J Magn Reson. 91:437-443
- 23 Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of  
24 isotopically enriched proteins. J Magn Reson 89:496-514
- 25 Kazimierczuk K, Zawadzka A, Koźmiński W, Zhukov I (2006) Random sampling of evolution time space and  
26 Fourier transform processing. J Biomol NMR 36:157-168
- 27 Kim S, Szyperski T (2003) GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral  
28 information. J Am Chem Soc 125:1385-1393
- 29 Kovacs H, Moskau D, Spraul M (2005) Cryogenically cooled probes - a leap in NMR technology. Prog NMR  
30 Spectrosc 46:131-155
- 31 Kupce E, Freeman R (2004) Projection-reconstruction technique for speeding up multidimensional NMR  
32 spectroscopy. J Am Chem Soc 126:6429-6440
- 33 Mäntylahti S, Hellman M, Permi P (2011) Extension of the HA-detection based approach: (HCA)CON(CA)H  
34 and (HCA)NCO(CA)H experiments for the main-chain assignment of intrinsically disordered proteins. J Biomol  
35 NMR 49:99-109
- 36 Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wüthrich K (1998)  
37 Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUPMB-IUPAB  
38 inter-union task group on the standardization of data bases of protein and nucleic acid structures determined  
39 by NMR spectroscopy. Eur J Biochem 256:1-15
- 40 Motackova V, Nováček J, Zawadzka-Kazimierczuk A, Kazimierczuk K, Židek L, Sanderová H, Krásný L,  
41 Koźmiński W, Sklenář V (2010) Strategy for complete NMR assignment of disordered proteins with highly  
42 repetitive sequences based on resolution-enhanced 5D experiments. J Biomol NMR 48:169-177

- 1 Narayanan RL, Duerr HN, Bilbow S, Biernat J, Mendelkow E, Zweckstetter M (2010) Automatic assignment of  
2 the intrinsically disordered protein Tau with 441-residues. *J Am Chem Soc* 132:11906-11907
- 3 Nováček J, Haba NY, Chill JH, Židek L, Sklenář V (2012) 4D Non-uniformly sampled HCBCACON and  $^1J_{NC\alpha}$ -  
4 selective HCBCANCO experiments for the sequential assignment and chemical shift analysis of intrinsically  
5 disordered proteins. *J Biomol NMR* 53:139-148
- 6 Nováček J, Zawadzka-Kazimierczuk A, Papoušková V, Židek L, Sanderová H, Krásný L, Koźmiński W, Sklenář  
7 V (2011) 5D  $^{13}C$ -detected experiments for backbone assignment of unstructured proteins with a very low signal  
8 dispersion. *J Biomol NMR* 50:1-11
- 9 Nováček J, Zawadzka-Kazimierczuk A, Papoušková V, Židek L, Sanderová H, Krásný L, Koźmiński W, Sklenář  
10 V (2011) 5D  $^{13}C$ -detected experiments for backbone assignment of unstructured proteins with a very low signal  
11 dispersion. *J Biomol NMR* 50:1-11
- 12 O'Hare B, Benesi AJ, Showalter SA (2009) Incorporating  $^1H$  chemical shift determination into  $^{13}C$ -direct  
13 detected spectroscopy of intrinsically disordered proteins in solution. *J Magn Reson.* 200:354-358
- 14 Orekhov VY, Ibragimov I, Billeter M (2003) Optimizing resolution in multidimensional NMR by three-way  
15 decomposition. *J Biomol NMR* 27:165-173
- 16 Piai A, Hošek T, Gonnelli L, Zawadzka-Kazimierczuk A, Koźmiński W, Brutscher B, Bermel W, Pierattelli R,  
17 Felli IC (2014) "CON-CON" assignment strategy for highly flexible intrinsically disordered proteins. *J Biomol*  
18 *NMR* 60:209-218
- 19 Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the  
20 structure determination of proteins in solution employing pulsed field gradients. *Progr NMR Spectrosc* 34:93-  
21 158
- 22 Shaka AJ, Barker PB, Freeman R (1985) Computer-optimized decoupling scheme for wideband applications  
23 and low-level operation. *J Magn Reson* 64:547-552
- 24 Tóth G, Gardai SJ, Zago W, Bertocini CW, Cremades N, Roy SL, Tambe MA, Rochet JC, Galvagnion C,  
25 Skibinski G, Finkbeiner S, Bova M, Regnstrom K, Chiou SS, Johnston J, Callaway K, Anderson JP, Jobling  
26 MF, Buell AK, Yednock TA, Knowles TP, Vendruscolo M, Christodoulou J, Dobson CM, Schenk D,  
27 McConlogue L (2014) Targeting the intrinsically disordered structural ensemble of  $\alpha$ -synuclein by small  
28 molecules as a potential therapeutic strategy for Parkinson's disease. *Plos ONE* 9:e87133
- 29 Uversky V, Oldfield CJ, Dunker AK (2008) Intrinsically disordered proteins in human diseases: introducing the  
30 D2 concept. *Annu Rev Biophys* 37:215-246
- 31 Uversky VN, Davé V, Iakoucheva LM, Malaney P, Metallo SJ, Pathak RR, Joerger AC (2014) Pathological  
32 unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem Rev* 114:6844-  
33 6879
- 34 van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer  
35 J, Jones DT, Kim PM, Kriwacki RW, Oldfield CJ, Pappu RV, Tompa P, Uversky VN, Wright PE, Babu MM  
36 (2014) Classification of intrinsically disordered regions and proteins. *Chem Rev* 114:6589-6631
- 37 Wright PE, Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol*  
38 *Cell Biol* 16:18-29
- 39 Yao X, Stefan B, Zweckstetter M (2014) A six-dimensional alpha proton detection-based APSY experiment  
40 for backbone assignment of intrinsically disordered proteins. *J Biomol NMR* 60: 231-240
- 41 Zawadzka-Kazimierczuk A, Kazimierczuk K, Koźmiński W (2010) A set of 4D NMR experiments of enhanced  
42 resolution for easy resonance assignment in proteins. *J Magn Reson* 202:109-116

1 Zawadzka-Kazimierczuk A, Koźmiński W, Sanderová H, Krásný L (2012) High dimensional and high resolution  
2 pulse sequences for backbone resonance assignment of intrinsically disordered proteins. J Biomol NMR  
3 52:329-337

4

5

6

## 1 **Figures**

2 **Figure 1.**  $^1\text{H}$ - $^{15}\text{N}$  HSQC (top) and  $^{13}\text{C}$ - $^{15}\text{N}$  CON-IPAP spectra (bottom), acquired at different  
3 temperatures (285.5 K, 295 K, 305 K, 315 K), showing the temperature dependence for  $\alpha$ -  
4 synuclein signals. Experiments were acquired at 16.4 T on  $\alpha$ -synuclein at pH 6.5.

5 **Figure 2.** A subset of the 2D projections acquired with the 5D (HCA)CONCACON  
6 experiment are reported to show the quality of the data and to provide an intuitive picture of  
7 how the APSY approach works with  $\text{C}'$  detected experiments on IDPs (Bermel et al., 2013).  
8 Orthogonal projections (A – D) and non-orthogonal ones (E – F) are shown. The spectra  
9 were recorded at 16.4 T on  $\alpha$ -synuclein at 315 K. The four orthogonal projections correspond  
10 to  $\text{C}'(i)\text{-N}(i+1)$ ,  $\text{C}'(i)\text{-N}(i)$ ,  $\text{C}'(i)\text{-C}'(i\pm 1)$ , and  $\text{C}^\alpha(i)\text{-C}'(i)$ . In the two non-orthogonal projections  
11 the y-coordinate units are arbitrary, as the frequencies in the indirect dimensions are a linear  
12 combination of frequencies, that is (E)  $\omega_2\sin(45^\circ) + \omega_4\cos(45^\circ)$  and (F)  $\omega_1\sin(55^\circ) +$   
13  $\omega_4\cos(55^\circ)$ .

14 **Figure 3.** As an example the assignment strategy using the peak lists resulting from the  
15 GAPRO analysis of the 5D (HCA)CONCACON projections is shown for residues 65-67 of  
16  $\alpha$ -synuclein (Bermel et al., 2013). Four five-dimensional “type I” correlations ( $\text{C}'_{i-1}$ ,  $\text{N}_i$ ,  $\text{C}^\alpha_i$ ,  
17  $\text{C}'_i$ ,  $\text{N}_{i+1}$ ) are shown. It is worth noting that the amide proton resonances of these residues  
18 were not detectable.

19

## 20 **Tables**

21 **Table 1.** Acquisition parameters for 5D  $^{\text{APSY}}$ -(HCA)CONCACON (Bermel et al., 2013) and  
22 4D  $^{\text{APSY}}$ -HCCCON experiments (Bermel et al., 2012).

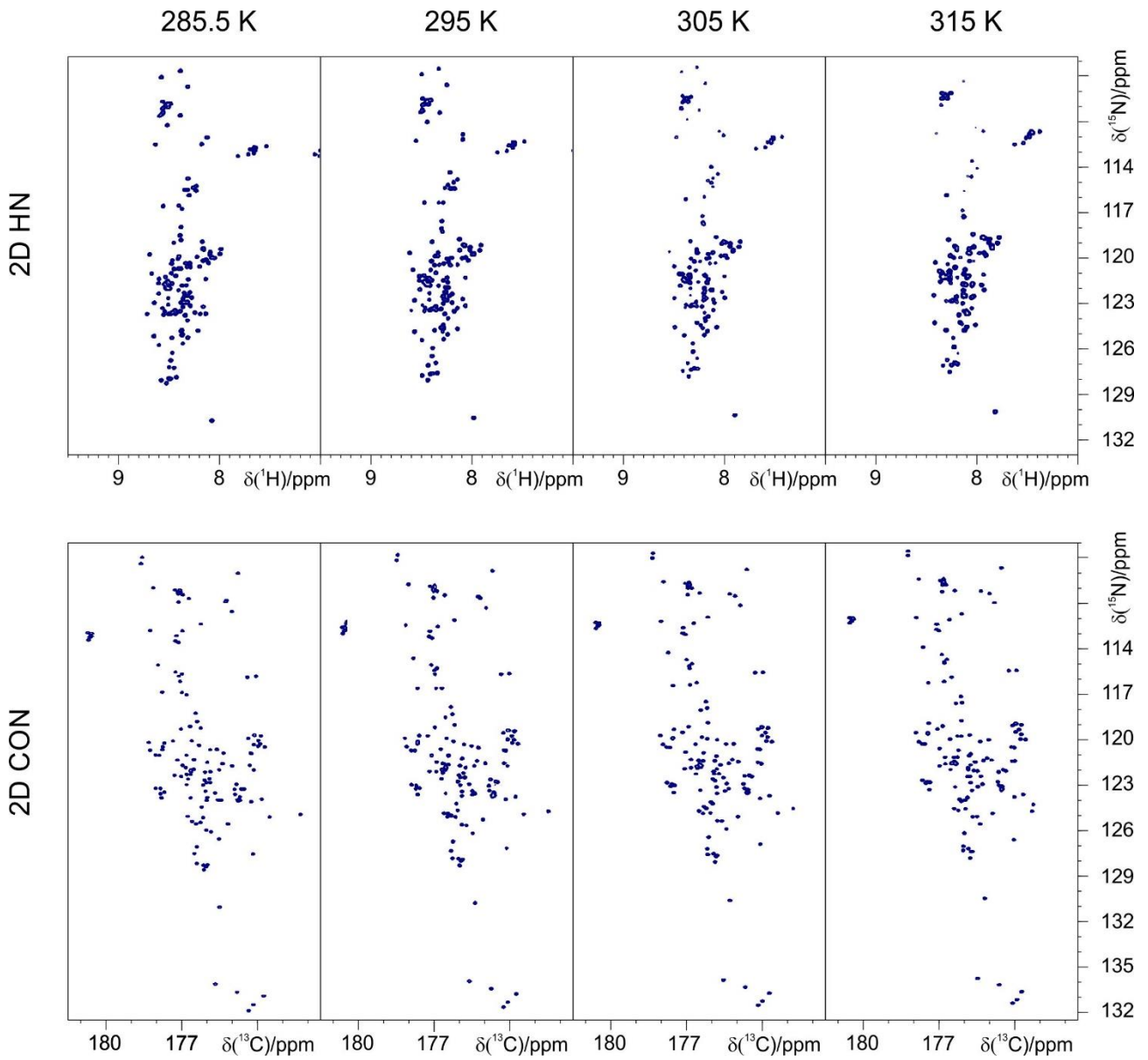
23 **Table 2.** Values of the projection angles  $\alpha$ ,  $\beta$ ,  $\gamma$ , and of the spectral widths in the dimensions  
24  $\omega_{1-4}$  used here for recording the 2D-projections of the 5D  $^{\text{APSY}}$ -(HCA)CONCACON  
25 experiment (Bermel et al., 2013). The resulting linear combinations of frequencies are given  
26 in the right column.

27 **Table 3.** Analysis of the 5D  $^{\text{APSY}}$ -(HCA)CONCACON processed considering a reduced  
28 number of projection angles as described in the text.

29

1 **Figure 1**

2



3

4

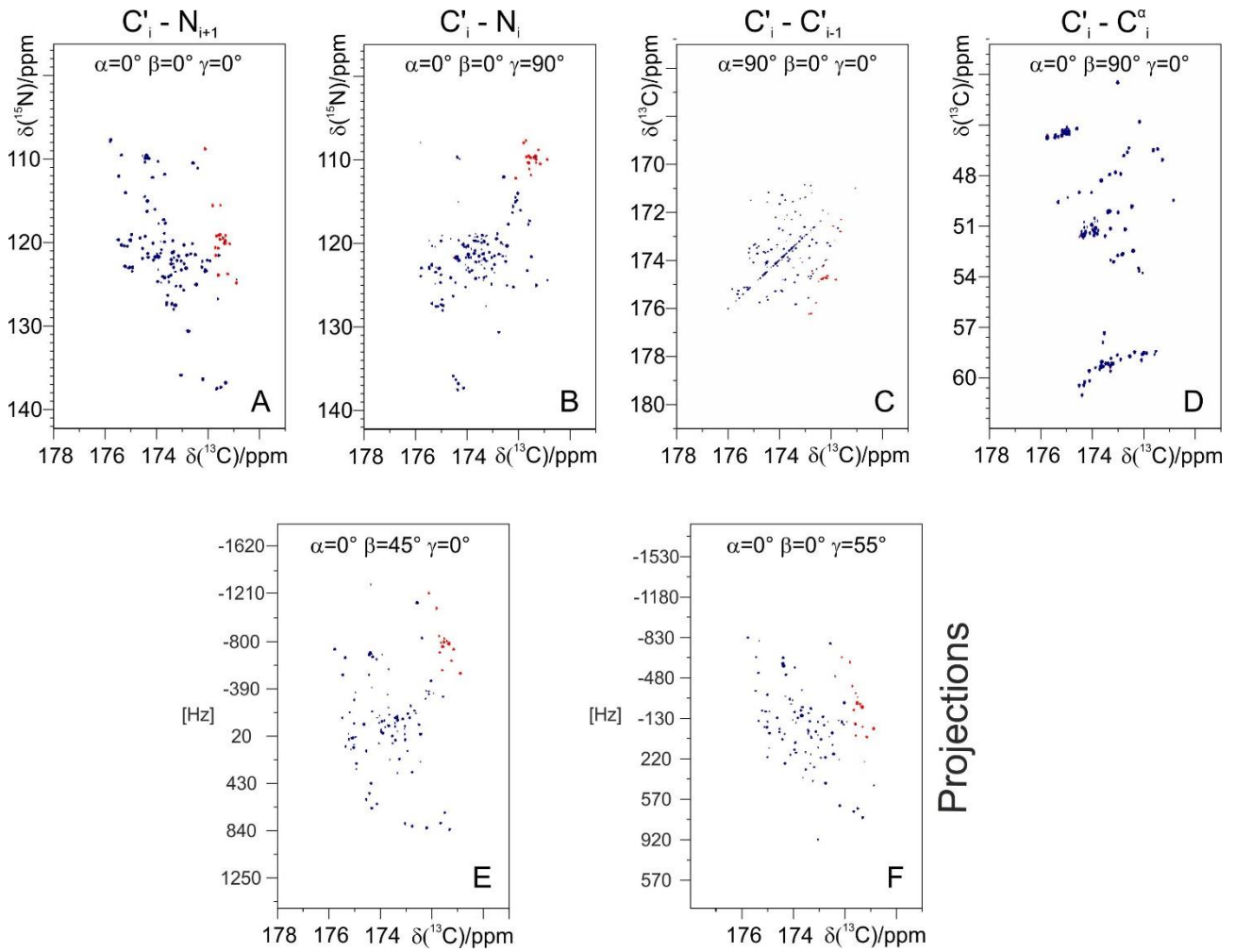
5

6

7

1  
2  
3  
4

**Figure 2**

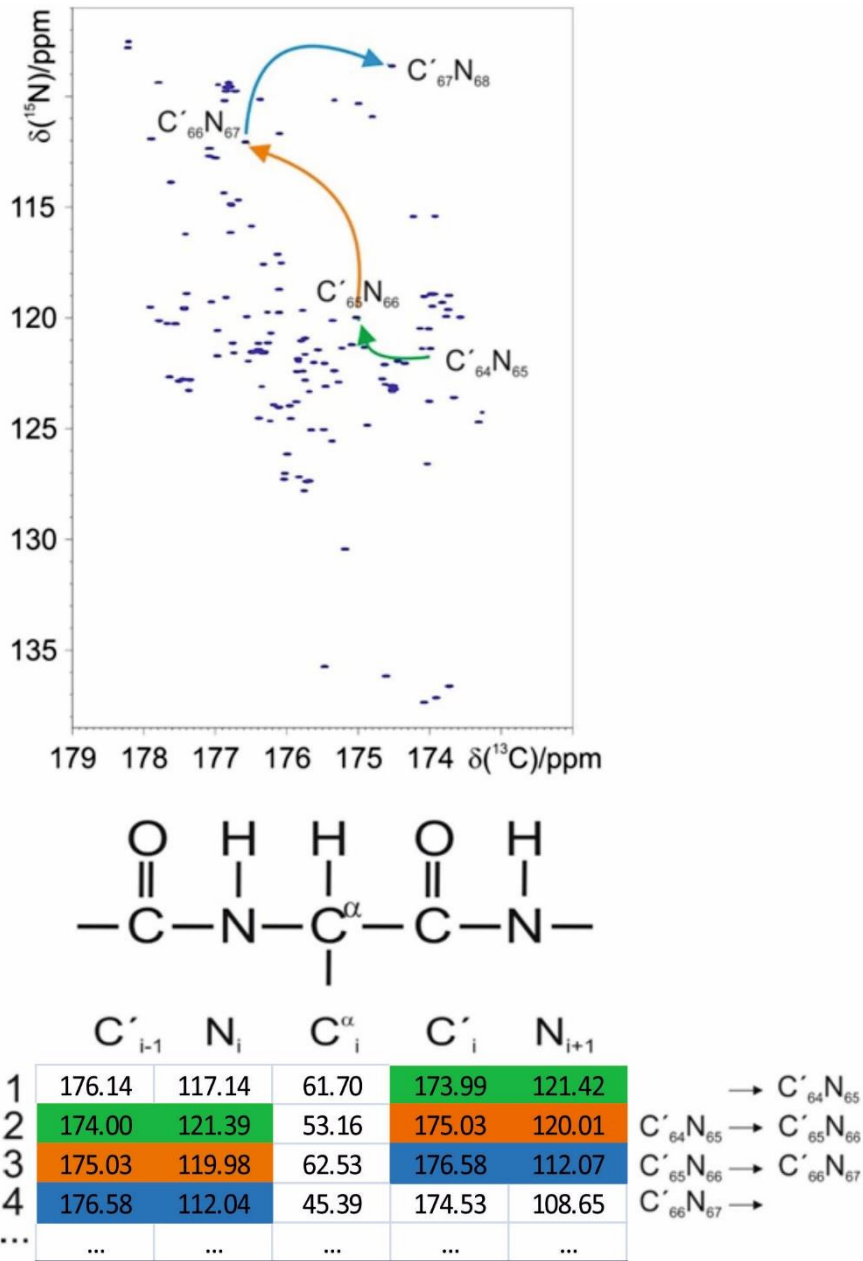


5

Projections

1  
2  
3

Figure 3



4  
5  
6  
7

**Table 1**

EXPERIMENT	Spectral width (Hz) & Maximal evolution times				NS	Inter scan delay (s)	Dimension of acquired data					Duration of the experiment	Number of projections
	F1	F2 / F4	F3	F5			F1	F2	F3	F4	F5		
5D (HCA)CONCACON	1695 ( <sup>13</sup> C')	2404 ( <sup>15</sup> N)	4505 ( <sup>13</sup> C <sup>α</sup> )	9090 ( <sup>13</sup> C')	8	0.8	256	256	256	256	2048	4 days 18 hours	64
4D HCCCON	F1	F2	F3	F4	4	0.8	512	512	512	2048	/	2 days 23 hours	41
	3840 ( <sup>1</sup> H)	10870 ( <sup>13</sup> C <sup>ali</sup> )	2404 ( <sup>15</sup> N)	9090 ( <sup>13</sup> C')									

**Table 2**

$\alpha$	$\beta$	$\gamma$	Spectral width (Hz)	Linear combination
0°	0°	0°	2404	$\omega_4$
0°	0°	90°	1695	$\omega_1$
0°	90°	0°	2404	$\omega_2$
90°	0°	0°	4505	$\omega_3$
$\pm 28^\circ$	0°	0°	4238	$\omega_4 \cos(28^\circ) \pm \omega_3 \sin(28^\circ)$
0°	$\pm 45^\circ$	0°	3400	$\omega_4 \cos(45^\circ) \pm \omega_2 \sin(45^\circ)$
0°	0°	$\pm 55^\circ$	2767	$\omega_4 \cos(55^\circ) \pm \omega_1 \sin(55^\circ)$
90°	$\pm 62^\circ$	0°	4238	$\omega_3 \cos(62^\circ) \pm \omega_2 \sin(62^\circ)$
90°	0°	$\pm 69^\circ$	3197	$\omega_3 \cos(69^\circ) \pm \omega_1 \sin(69^\circ)$
0°	90°	$\pm 55^\circ$	2767	$\omega_2 \cos(55^\circ) \pm \omega_1 \sin(55^\circ)$
$\pm 28^\circ$	$\pm 41^\circ$	0	4775	$\omega_4 \cos(28^\circ) \cos(41^\circ) \pm \omega_2 \sin(41^\circ) \pm \omega_3 \sin(28^\circ) \cos(41^\circ)$
$\pm 28^\circ$	0°	$\pm 51^\circ$	3984	$\omega_4 \cos(28^\circ) \cos(51^\circ) \pm \omega_1 \sin(51^\circ) \pm \omega_3 \sin(28^\circ) \cos(51^\circ)$
0°	$\pm 45^\circ$	$\pm 45^\circ$	3603	$\omega_4 \cos(45^\circ) \cos(45^\circ) \pm \omega_1 \sin(45^\circ) \pm \omega_2 \sin(45^\circ) \cos(45^\circ)$
90°	$\pm 62^\circ$	$\pm 51^\circ$	3984	$\omega_3 \cos(62^\circ) \cos(51^\circ) \pm \omega_1 \sin(51^\circ) \pm \omega_2 \sin(62^\circ) \cos(51^\circ)$
$\pm 15^\circ$	0°	0°	3488	$\omega_4 \cos(15^\circ) \pm \omega_3 \sin(15^\circ)$
0°	$\pm 27^\circ$	0°	3233	$\omega_4 \cos(27^\circ) \pm \omega_2 \sin(27^\circ)$
0°	0°	$\pm 35^\circ$	2941	$\omega_4 \cos(35^\circ) \pm \omega_1 \sin(35^\circ)$
90°	$\pm 43^\circ$	0°	4934	$\omega_3 \cos(43^\circ) \pm \omega_2 \cos(43^\circ)$
90°	0°	$\pm 53^\circ$	4065	$\omega_3 \cos(53^\circ) \pm \omega_1 \sin(53^\circ)$
0°	90°	$\pm 35^\circ$	2941	$\omega_2 \cos(35^\circ) \pm \omega_1 \sin(35^\circ)$
$\pm 47^\circ$	0°	0°	4934	$\omega_4 \cos(47^\circ) \pm \omega_3 \sin(47^\circ)$
0°	$\pm 63^\circ$	0°	3233	$\omega_4 \cos(63^\circ) \pm \omega_2 \sin(63^\circ)$
0°	0°	$\pm 71^\circ$	2385	$\omega_4 \cos(71^\circ) \pm \omega_1 \sin(71^\circ)$
90°	$\pm 75^\circ$	0°	3488	$\omega_3 \cos(75^\circ) \pm \omega_2 \cos(75^\circ)$
90°	0°	$\pm 79^\circ$	2523	$\omega_3 \cos(79^\circ) \pm \omega_1 \sin(79^\circ)$
0°	90°	$\pm 71^\circ$	2385	$\omega_2 \cos(71^\circ) \pm \omega_1 \sin(71^\circ)$
$\pm 15^\circ$	0°	0°	3488	$\omega_4 \cos(15^\circ) \pm \omega_3 \sin(15^\circ)$
0°	$\pm 20^\circ$	0°	3081	$\omega_4 \cos(20^\circ) \pm \omega_2 \sin(20^\circ)$
$\pm 70^\circ$	0°	0°	5056	$\omega_4 \cos(70^\circ) \pm \omega_3 \sin(70^\circ)$
0°	$\pm 70^\circ$	0°	3081	$\omega_4 \cos(70^\circ) \pm \omega_2 \sin(70^\circ)$

**Table 3**

Angle sets	Projections	Experimental time	Identified correlations (Type I)
30	64	4d, 18h	131
26	56	4d, 5h	131
20	44	3d, 9h	129
10	16	1d, 2h	116

# **$^{13}\text{C}$ APSY-NMR for sequential assignment of intrinsically disordered proteins**

Maria Grazia Murralli<sup>1</sup>, Marco Schiavina<sup>1</sup>, Valerio Sainati<sup>1</sup>, Wolfgang Bermel<sup>2</sup>, Roberta Pierattelli<sup>1,3</sup>✉ and Isabella C. Felli<sup>1,3</sup>✉

## **Supporting material**

**Figure S1.** Pulse scheme for the 4D CT-HCCCON. The following phase cycling was employed:  $\phi_1 = x, -x$ ;  $\phi_2 = 2(y), 2(-y)$ ;  $\phi_3 = 4(x), 4(-x)$ ;  $\phi_4 = 16(x), 16(-x)$ ;  $\phi_5 = 8(x), 8(-x)$ ,  $\phi_{rec} = x, 2(-x), x, (-x), 2(x), (-x), (-x), 2(x), (-x), x, 2(-x), x$ . The quadrature detection was achieved through the States-TPPI approach by incrementing the phase of the  $\pi/2$  pulse prior to the evolution period;  $C'$  homonuclear decoupling was achieved using IPAP scheme. The length of the delays were:  $\delta = 0.475$  ms,  $\Delta = 56$  ms,  $\Delta_1 = 9$  ms,  $\Delta_2 = 25$ ms -  $p_{12}$ ,  $\epsilon = t_3(0) + p_8$ .  $t_a$ ,  $t_b$  and  $t_c$  were used to achieve the semi-constant time mode for the  $^1H$  indirect dimension. The  $^1H$  carrier was placed in the middle of the  $^1H$  aliphatic region (2.5 ppm) during the evolution block while it was set at 4.7 ppm in the remaining parts of the pulse sequence.

**Figure S2.** Examples of 2D projections acquired with the 4D CT-HCCCON experiment to show the quality of the data. Three orthogonal projections (A – C) and a non-orthogonal one (D) are shown. The spectra were recorded at 700 MHz on  $\alpha$ -synuclein at 315 K. The three orthogonal projections correspond to  $C'(i)-N(i+1)$ ,  $C'(i)-Hali(i)$ ,  $C'(i)-Cali(i)$ . The non-orthogonal projection has, as indirect dimension, the combination of  $\omega_1 \sin(-32^\circ) + \omega_3 \cos(-32^\circ)$ .

Figure S1.

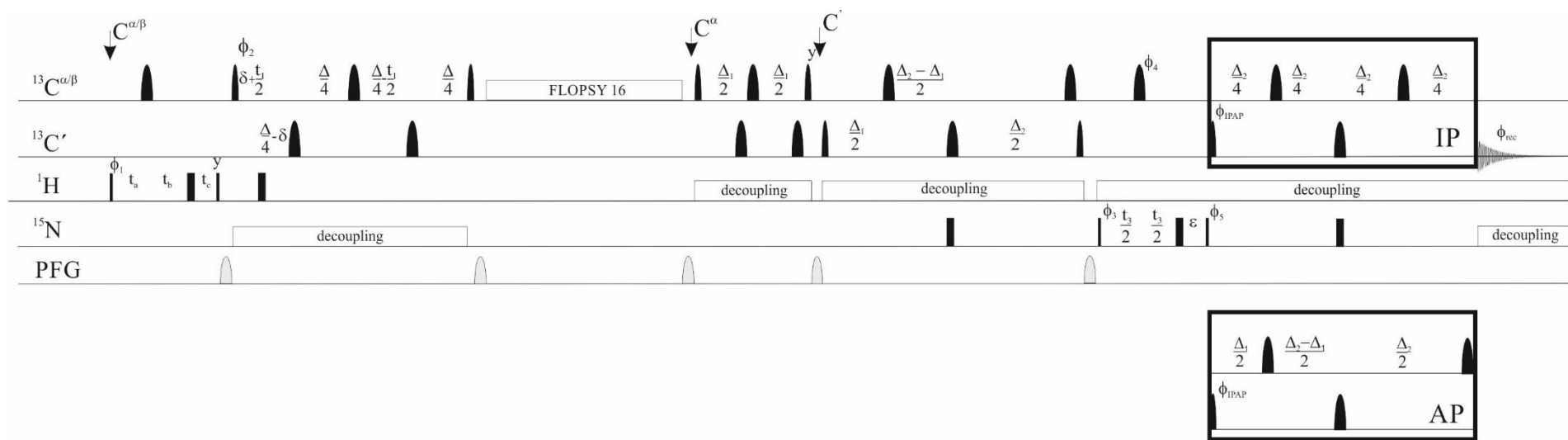


Figure S2

