Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2019

Supporting Information

Rapid NMR assignments of Intrinsically Disordered Proteins using two-dimensional ^{13}C-detection based experiments

Sujeesh Sukumaran [a,b], Shahid Malik [a,b], Shankararama Sharma R [b], Kousik Chandra *[b], and Hanudatta Atreya *[b]

a. Solid State and Structural Chemistry Unit (SSCU), Indian Institute Of Science, Bangalore – 560 012 India

b. NMR Research Centre, Indian Institute of Science, Bangalore - 560 012, India

Corresponding Authors

Hanudatta Atreya

Email: hsatreya@iisc.ac.in

Kousik Chandra

Email: kousikchandra@gmail.com

Protein Expression and purification

α-synuclein in humans is encoded by the SNCA gene. It is abundant in the brain while smaller amounts are found in the heart, muscles, and other tissues. In the brain, α-synuclein is found mainly at the tips of nerve cells (neurons) in specialized structures called presynaptic terminals. The amino acid sequence of the protein is shown below

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATV
AEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGI
LEDMPVDPDNEAYEMPSEEGYQDYEPEA

Transformation

We have transformed the α-synuclein gene into *E.coli* BL21 DE3 competent cells by heat shock method at 42 °C for 90 seconds.

Cell Culture and isotopic labeling

The transformed cells were streaked on a LB (Luria Bertani) -Agar plate containing 100 µg/ml of Ampicillin and incubated at 37 °C for overnight. A well isolated single colony from the plate was inoculated to a primary culture consists of 5 ml of LB medium and 100 µg/ml of Ampicillin and incubated at 37 °C for overnight. NMR characterization of the protein requires uniform isotopic labeling of the protein. The incorporation of NMR active isotopes i.e. ^{13}C and ^{15}N into protein was achieved by growing the transformed BL21 (DE3) cells from the primary culture in M9 minimal medium containing $^{13}C_6$-glucose and $^{15}NH_4Cl$ as the sole source of carbon and nitrogen. For selective unlabeling, 1.0 g/l each of desired unlabeled amino acids: Thr, Gly, Ser and Ala (stock solution of 1.0 g in 50 ml of H_2O was prepared and filter sterilized) was added to the growth medium. Cells from the primary culture was inoculated to 1000 ml minimal media containing 100 µg/ml of Ampicillin and grown at 37 °C until the OD_{600} reached 0.8.

Over-expression and purification

Over-expression of the protein was achieved by inducing the cells with 1 mM Isopropyl β-D-Thioogalactoside (IPTG) for 5 hours at 37 °C. Cells were harvested by centrifugation at 5000 rpm for 10 minutes. Cell lysis was performed by suspending cells in lysis buffer (Composition: 50 mM Tris (tris(hydroxymethyl)aminomethane), at pH 8 supplemented with 10mM EDTA and 150 mM NaCl) followed by sonication in 10 seconds cycles, with an intervening period of 20 seconds, for a total time of 30 minutes. After sonication the cell lysate was centrifuged at 12,000 rpm for 60 minutes at 4 °C to separate insoluble debris from soluble fraction containing the protein. The supernatant was collected and 10% streptomycin and glacial acetic acid was added to solution. This mixture was kept in ice for 60 minutes and then centrifuged at 12,000 rpm for 1 hour. The supernatant after this centrifugation was collected and equal volume of saturated ammonium sulphate was added to it. This mixture was incubated at 4 °C overnight with constant stirring. This mixture was then centrifuged at 12000 rpm for 60 minutes at 4 °C. The pellet of the

centrifuged mixture was then used for purification. This pellet was resuspended in 50 % saturated ammonium sulphate (1:1 mixture with water) and then incubated in ice for 60 minutes and then centrifuged at 12000 rpm, 4 °C, for 60 minutes. Later the pellet was isolated from this step and added 100 mM ammonium acetate and 20 ml of ethanol to it. This mixture was kept in ice for 60 minutes and then centrifuged at 12000 rpm, 4 °C, for 60 minutes. Later 40 ml of ethanol was added to the pellet, kept in ice for 60 minutes and then centrifuged at 12000 rpm, 4 °C, for 60 minutes. Then pellet from this stage was resuspended in 6 ml of ammonium acetate kept in ice for 60 minutes and then centrifuged at 12000 rpm, 4 °C, for 60 minutes. The final pellet contains the pure protein. The purity of the sample was confirmed by loading it on 15% SDS Page resolving gel.

3.2.2 Nuclear Magnetic Resonance Spectroscopy

Sample preparation

The purified protein (0.4 mM) was solubilized in 50 mM Phosphate buffer (pH 7) containing 50 mM NaCl, 0.01% sodium azide and 5% D_2O (for locking). All NMR data were recorded at 283 K on a Bruker Avance III 800 MHz spectrometer equipped with a cryogenically cooled triple resonance probe. All experiments were recorded in 5 mm sample tube with sample volume of 0.5 ml.

NMR experiments

Table 2.1

Experiment	Nucleus	FID Size	# Scans	Spectral Width (ppm)	Offset (ppm)	Experimental Time
HSQC	^{1}H	2048	4	12	4.7	22 minutes
	^{15}N	256		34	117	
hNcoCA	^{15}N	480	32	32	119	5 Hours
	^{13}C	1024		40	54	
hNCA	^{15}N	512	80	34	119	13 hours 9 minutes
	^{13}C	1024		40	54	

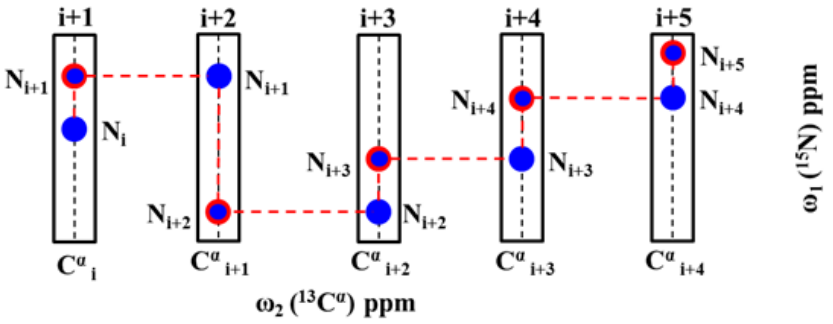
The 2D [^{15}N, ^{1}H] HSQC is not required for resonance assignments. However, it is recorded to check the stability of the sample. Note that the same measurement time is used for both uniformly labeled and selectively unlabeled sample.

Supporting information Figure S1:

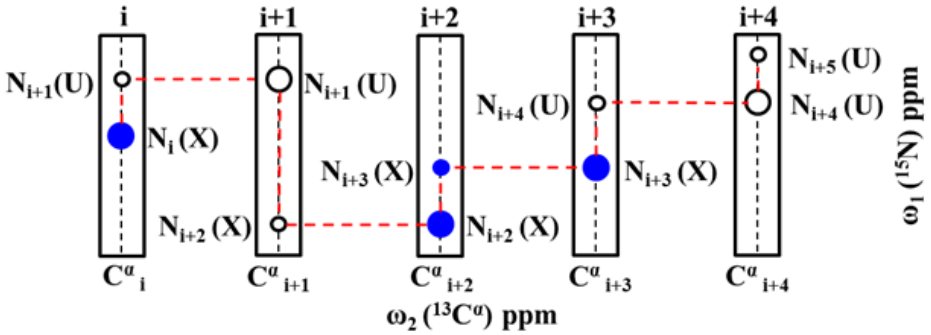
Step 1: Consider protein primary sequence and translate it to a sequence of X (labeled) and U (unlabeled)

e.g. Sequence – **EFGGKKLELETSVSDLNVHGF**
Unlabeled residues: **T, G and S**
Translated sequence - **XXUUXXXXXXUUXUXXUXXUX**

Step 2: Analyze 2D hNcoCa (Red) and 2D hNCa (Blue) to obtain sequential connectivity.



Step 3: Identify the pattern of X and U in a connected stretch using 2D hNCa spectrum recorded on selectively unlabeled sample. The schematic shows the peaks pattern for XUXXUU



Step 4: If the identified pattern is unique in the primary sequence, we can directly assign the resonances

G V A T V A E K T K E Q V T N V G G A V V T G V T A V A Q
U X U U X U X X U X X X U X X U U U X X U U X U U X U X

Fig. S1: Complete assignment strategy

Supporting information Figure S2:

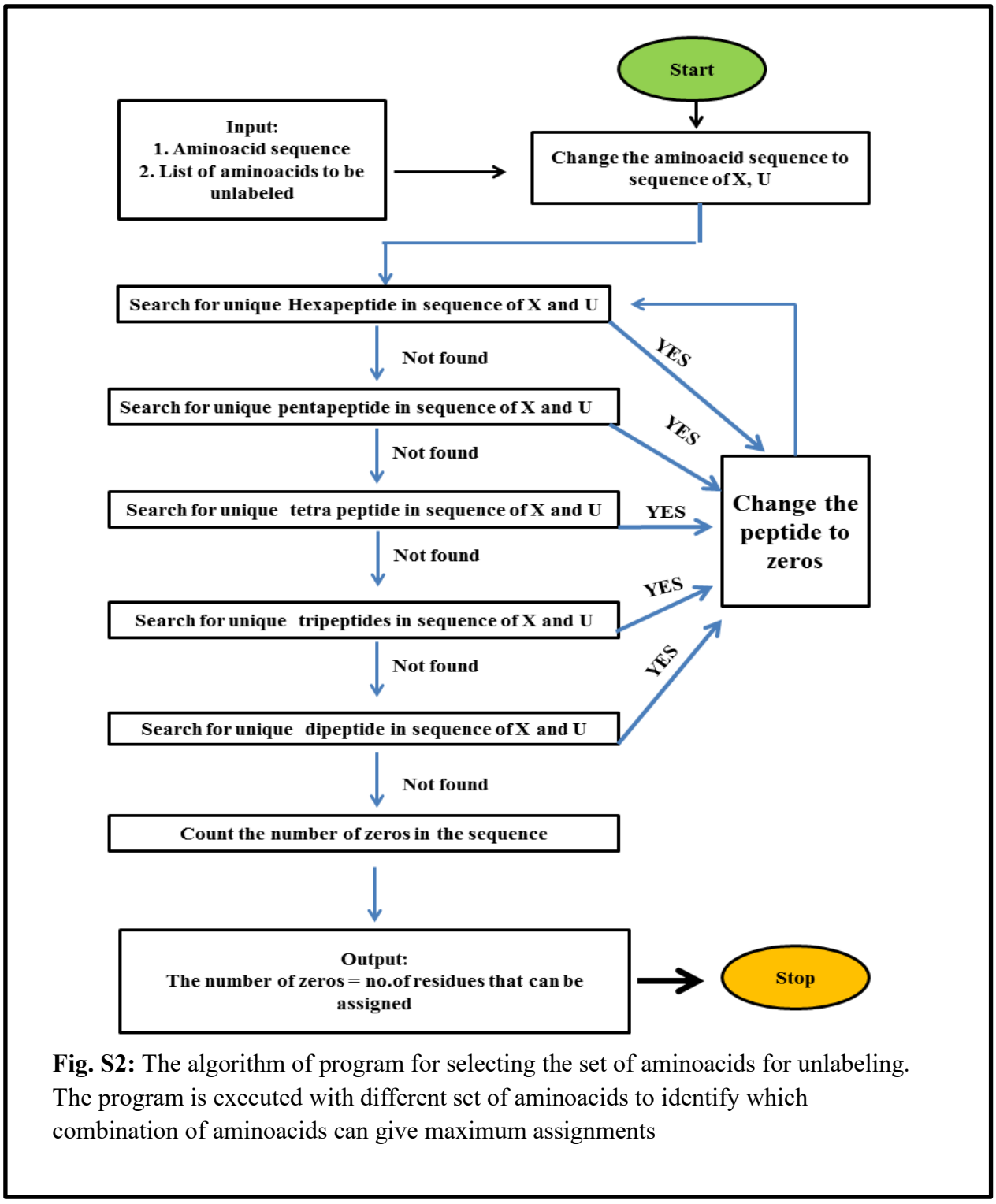


Fig. S2: The algorithm of program for selecting the set of aminoacids for unlabeling. The program is executed with different set of aminoacids to identify which combination of aminoacids can give maximum assignments

Supporting information Figure S3:

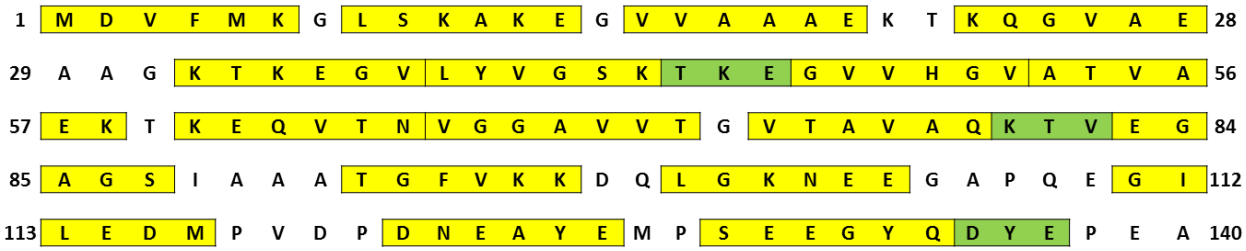


Fig. S3: Stretches of amino acid residues along the polypeptide chain of α-synuclein which could be assigned using the strategy depicted in Fig. S1. The residues shown in Yellow block are hexapeptide stretches which are initially assigned, followed by shorter stretches shown in the Green blocks.