

**Protein Structure Characterization by Solid-State
NMR: Structural Comparison of Mouse and Human α -
Synuclein Fibrils, Sparse ^{13}C Labeling Schemes, and
Stereospecific Assignment of Val and Leu Prochiral
Methyl Groups**

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

vorgelegt von

Guohua Lv

aus Jining, China

Göttingen, 2013

Mitglied des Betreuungsausschusses (Referent): Dr. Adam Lange
NMR-basierte Strukturbiologie, Max-Planck-Institut für biophysikalische Chemie

Mitglied des Betreuungsausschusses (Koreferent): Prof. Dr. Kai Tittmann
Abteilung für Bioanalytik, Georg-August-Universität Göttingen

Mitglied des Betreuungsausschusses: Prof. Dr. Berend de Groot
Computergestützte biomolekulare Dynamik, Max-Planck-Institut für biophysikalische Chemie

Tag der mündlichen Prüfung:

Affidavit

I hereby declare that this dissertation has been written independently and with no other sources and aids than quoted.

.....

Guohua Lv

Table of Content

Affidavit	III
Abbreviations	VII
Related publications	IX
Abstract	X
Part I Structural Comparison of Mouse and Human α-Synuclein Amyloid Fibrils and Sparse ^{13}C Labeling Schemes	1
1 Introduction	2
1.1 Solid-state nuclear magnetic resonance	2
1.2 Sparse labeling schemes	7
1.3 Parkinson's disease and α -synuclein fibrils.....	11
2 Materials and methods	19
2.1 Sample preparation and morphological characterization.....	19
2.1.1 α -Synuclein protein expression in <i>E. coli</i> and purification	19
2.1.2 Fibrillization <i>in vitro</i>	20
2.1.3 Electron microscopy	20
2.2 Solid-state NMR experiments and analysis	21
2.2.1 Experiments on uniformly [^{13}C]glucose labeled mouse α -synuclein fibrils	
21	
2.2.1.1 2D ^{13}C - ^{13}C and ^{15}N - ^{15}N correlation experiments.....	21
2.2.1.2 2D/3D ^{15}N - ^{13}C correlation experiments.....	22
2.2.1.3 The indirect detection of ^1H - ^1H correlations	22
2.2.2 Experiments on [$2\text{-}^{13}\text{C}$]glucose labeled mouse α -synuclein fibrils	22
2.2.3 Experiments on [$1\text{-}^{13}\text{C}$]glucose labeled mouse α -synuclein fibrils	23
2.2.4 Experiment on mixed 1:1 ^{13}C : ^{15}N labeled mouse α -synuclein fibrils	23
3 Results and discussion	25
3.1 Morphological characterization by electron microscopy and initial comparison to human α -synuclein fibrils.....	25

3.2	Sequential resonance assignment using uniformly [¹³ C]glucose labeled mouse α -synuclein fibrils	28
3.3	Sparse labeling schemes: [2- ¹³ C]glucose and [1- ¹³ C]glucose	33
3.3.1	Labeling pattern of [2- ¹³ C]Glc labeling scheme	33
3.3.2	Resolution enhancement and spectrum simplification using [2- ¹³ C]Glc labeling scheme.....	41
3.3.3	Labeling pattern of [1- ¹³ C]glucose labeling scheme	50
3.3.4	Resolution enhancement and spectrum simplification using [1- ¹³ C]glucose labeling scheme.....	55
3.4	Secondary structure analysis of the fibril core of mouse α -synuclein fibrils	62
3.5	Structural comparison of mouse and human α -synuclein fibrils.....	66
3.6	Supra-molecular structure of mouse α -synuclein fibrils elucidated by means of mixed 1:1 ¹³ C: ¹⁵ N labeled mouse α -synuclein fibrils	69
4	Conclusions	71
Part II	Stereospecific Assignment of Val and Leu Prochiral Methyl Groups	72
1	Introduction.....	73
2	Experiments and methods.....	76
2.1	Sample preparation	76
2.2	Solid-state NMR experiments and data processing	76
2.3	Solution NMR experiments and data processing.....	77
3	Results and discussion	78
3.1	Key metabolic pathways involved in the formation of pyruvate.....	78
3.2	Stereoselective biosynthesis of Val and Leu	81
3.3	Stereospecific assignment of Val and Leu in T3SS PrgI needles and ubiquitin	83
4	Conclusions	88
Part III	Summary	89
Part IV	Appendix.....	91
	Appendix A	92
	Appendix B.....	107
	References.....	110

Acknowledgements	121
Curriculum Vitae	123

Abbreviations

[1- ¹³ C]Glc-labeled	[1- ¹³ C]glucose labeled
[2- ¹³ C]Glc-labeled	[2- ¹³ C]glucose labeled
[U- ¹³ C]Glc-labeled	uniformly [¹³ C]glucose labeled
[M- ¹³ C/ ¹⁵ N]-labeled	mixed 1:1 ¹³ C: ¹⁵ N labeled
3PG	3-phosphoglycerate
AFM	atomic force microscopy
AKG	α-ketoglutarate
αS	α-synuclein
CoA	coenzyme A
CP	cross-polarization
CSA	chemical-shift anisotropy
CW	continuous wave
DAP	L, L-α,ε-diaminopimelate
DARR	dipolar assisted rotational resonance
DHAP	dihydroxyacetone phosphate
DREAM	dipolar recoupling enhancement through amplitude modulation
DSS	sodium salt of 2,2-dimethyl-2-silapentane-5-sulphonic acid
E4P	erythrose-4-phosphate
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediamine tetraacetic acid
EM	electron microscopy
ESI-MS	electrospray ionization mass spectrometry
F6P	fructose-6-phosphate
Far-UV CD	Far-UV Circular Dichroism spectroscopy
FTIR	Fourier transform infrared spectroscopy
γ	gyromagnetic ratio

G6P	glucose-6-phosphate
GAP	glyceraldehyde-3-phosphate
h α S	human α -synuclein
H/D	hydrogen/deuterium
INEPT	insensitive nuclei enhanced by polarization transfer
LBs	Lewy bodies
LN _s	Lewy neuritis
MAS	magic-angle spinning
m α S	mouse α -synuclein
OA	oxaloacetate
PAGE	polyacrylamide gel electrophoresis
PD	Parkinson's disease
PDS _D	proton-driven spin diffusion
PEP	phosphoenolpyruvate
PMSF	phenylmethylsulfonyl fluoride
PPP	pentose phosphate pathway
R5P	ribose-5-phosphate
Ru5P	ribulose-5-phosphate
r.f.	Radio-frequency
S7P	sedoheptulose-7-phosphate
SPC5	supercycled POST-C5
SPINAL	small phase incremental alternation decoupling
ssNMR	solid-state nuclear magnetic resonance
TCA cycle	tricarboxylic acid cycle
ThT fluorescence	thioflavin T fluorescence
TPP	thiamine pyrophosphate
Xu5P	xylulose-5-phosphate

Related publications

The work presented in this thesis is based on the following publications. I am very grateful to all co-authors of these papers, as well as the people acknowledged in them for the great cooperations.

- **Guohua Lv**, Hannes Klaus Faßhuber, Antoine Loquet, Jean-Philippe Demers, Vinesh Vijayan, Karin Giller, Stefan Becker, and Adam Lange* A Straightforward Method for Stereospecific Assignment of Val and Leu Prochiral Methyl Groups by Solid-State NMR: Scrambling in the [2-¹³C]Glucose Labeling Scheme. *J. Magn. Reson.* 2013 Mar 29;228:45-49. (Cover Article)
- **Guohua Lv**, Ashutosh Kumar, Karin Giller, Maria L. Orcellet, Dietmar Riedel, Claudio O. Fernández, Stefan Becker, and Adam Lange*. Structural Comparison of Mouse and Human α -Synuclein Amyloid Fibrils by Solid-State NMR. *J. Mol. Biol.* 2012 Jun 29;420(1-2):99-111.
- Antoine Loquet, **Guohua Lv**, Karin Giller, Stefan Becker, and Adam Lange*. ¹³C Spin Dilution for Simplified and Complete Solid-State NMR Resonance Assignment of Insoluble Biological Assemblies. *J. Am. Chem. Soc.* 2011 Apr 6;133(13):4722-4725.

Abstract

This thesis describes protein structure characterization by solid-state nuclear magnetic resonance (ssNMR), including structural investigation of mouse α -synuclein ($m\alpha S$) fibrils and the comparison to human α -synuclein ($h\alpha S$) fibrils, sparse ^{13}C labeling schemes based on $[1-^{13}\text{C}]\text{Glucose}$ ($[1-^{13}\text{C}]\text{Glc}$) and $[2-^{13}\text{C}]\text{Glucose}$ ($[2-^{13}\text{C}]\text{Glc}$), and stereospecific assignment of Val and Leu prochiral methyl groups by the $[2-^{13}\text{C}]\text{Glc}$ labeling scheme.

Fibrillar αS is the major component of Lewy bodies, the pathological hallmark of Parkinson's disease. $m\alpha S$ aggregates much faster than $h\alpha S$, although $m\alpha S$ differs from $h\alpha S$ at only seven positions in its primary sequence. Currently, little is known about the site-specific structural differences between $m\alpha S$ and $h\alpha S$ fibrils. In Part I of this thesis, we applied state-of-the-art ssNMR methods to structurally characterize $m\alpha S$ fibrils. The assignment strategy employed a set of high-resolution 2D and 3D ssNMR spectra recorded on uniformly $[^{13}\text{C}]\text{Glc}$ -, $[1-^{13}\text{C}]\text{Glc}$ -, and $[2-^{13}\text{C}]\text{Glc}$ -labeled $m\alpha S$ fibrils. Proteins produced with $[1-^{13}\text{C}]\text{Glc}$ or $[2-^{13}\text{C}]\text{Glc}$ are very sparsely labeled and the resulting 2D ssNMR spectra exhibit smaller linewidths and contain a reduced number of cross peaks. This allows for an accelerated and straightforward resonance assignment using a limited set of simple 2D experiments. The approach is successfully demonstrated with the *de novo* assignment of $m\alpha S$ fibrils.

The unambiguous stereospecific assignment of the prochiral methyl groups in Val and Leu plays an important role in the structural investigation of proteins by NMR. In Part II of this thesis, we present a straightforward method for the stereospecific ssNMR assignment of the prochiral methyl groups of Val and Leu using $[2-^{13}\text{C}]\text{Glc}$ as the sole carbon source for the biosynthesis of amino acids. The approach is fundamentally based on the stereoselective biosynthetic pathway of Val and Leu, and the co-presence of $[2-^{13}\text{C}]\text{pyruvate}$ produced mainly by glycolysis and $[3-^{13}\text{C}]/[1,3-^{13}\text{C}]\text{pyruvate}$ most probably formed through scrambling in the pentose phosphate pathway starting from $[2-^{13}\text{C}]\text{Glc}$. As a consequence, the isotope spin pairs of $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma 2$ and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma 1$ in Val, and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta 2$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta 1$ in Leu are obtained. The approach is successfully demonstrated

with the stereospecific assignment of the methyl groups of Val and Leu of type 3 secretion system PrgI needles and of microcrystalline ubiquitin.

**Part I Structural Comparison of Mouse and Human
 α -Synuclein Amyloid Fibrils and Sparse ^{13}C labeling
Schemes**

1 Introduction

1.1 Solid-state nuclear magnetic resonance

Due to recent developments in magic-angle spinning (MAS) solid-state nuclear magnetic resonance (ssNMR) methodology [1; 2; 3; 4], isotope labeling schemes [5; 6; 7], structure calculation protocols [8; 9], and access to high-field spectrometers, ssNMR has emerged as a powerful tool for the structural characterization of noncrystalline and insoluble proteins [3], including membrane proteins [10; 11; 12; 13; 14; 15], oligomeric assemblies [16; 17; 18; 19; 20; 21], and amyloid fibrils [22; 23; 24; 25; 26; 27; 28; 29]. One major advantage of ssNMR spectroscopy over other biophysical techniques used for structural investigation of proteins, e.g., X-ray crystallography and solution NMR, is that no need for the protein to be crystallized or to be soluble. However, for biomolecules uniformly labeled with ^{13}C , ssNMR often exhibits broadened signals due to the sample inhomogeneity and the influence of orientation-dependent anisotropic interactions that are also present, but rarely observed in solution NMR due to rapid molecular tumbling. The three primary interactions are the heteronuclear and homonuclear dipolar couplings (through-space), the chemical-shift anisotropy (CSA), and the scalar J couplings (through-bond).

a) Heteronuclear dipolar coupling and homonuclear dipolar coupling, arising from an interaction between the nuclear magnetic moments of two nuclear spins (e.g., ^1H - ^{13}C for heteronuclear, ^{13}C - ^{13}C for homonuclear) through space [30]. The magnitude of the dipolar coupling (Eq. (1) and (2)) is proportional to the product of the gyromagnetic ratio (i.e., γ) of the nuclei of interest, and inversely proportional to the cube of the inter-nuclear distance (i.e., r^{-3}). It also depends on the orientation (i.e., the term of $(1-3\cos^2\theta)$), where θ is the angle between the inter-nuclear vector and the static external magnetic field B_0).

$$H_D = d(\theta) \cdot [3 \cdot S_z I_z - S \cdot I] \quad (1)$$

S_z and I_z are the Cartesian Z components of spin operators S and I.

$$d(\theta) = \frac{\mu_0}{4\pi} \frac{\gamma_s \gamma_I \hbar}{r^3} \frac{1-3\cos^2(\theta)}{2} \quad (2)$$

where r is the inter-nuclear distance, γ is the gyromagnetic ratio, μ_0 and \hbar are the permeability of free space and Planck's constant [31], respectively.

b) CSA, resulting from interactions of the nuclear spins with surrounding electrons, i.e., when an external magnetic field is applied to an atom, not only are the nuclear spins perturbed, but the surrounding electrons are also affected since they have magnetic moments too. The external field induces circulating currents of electrons that in turn produce small magnetic fields, which either add to or subtract from the external field felt by the nucleus [30]. Consequently, the effective magnetic field experienced by the nucleus is thus altered, as is its resonance frequency (i.e., chemical shift).

c) Scalar J coupling, resulting from indirect spin-spin interactions between two nuclei mediated by the electrons of the chemical bond [31]. The scalar J coupling exhibits no spatial orientation dependence, and remains in the presence of fast molecular tumbling [30]. Thereby, the J coupling can be exploited for magnetization transfer in both ssNMR and solution NMR spectroscopy. Typically, the one-bond J coupling constants for $^1J_{C\alpha,C\beta}$ and $^1J_{C\alpha,C'}$ are 35 Hz and 55 Hz [32], respectively.

Achieving high resolution ssNMR spectra requires the suppression of the above mentioned interactions. In order to average out the orientation-dependent anisotropic interactions, i.e., dipolar couplings and CSA, the most efficient and robust technique in ssNMR, MAS [33; 34] is used. MAS exploits the orientation dependence of the anisotropic interactions which vary with the term of $(1-3\cos^2\theta)$. Spinning of a solid-phase sample at an angle of 54.7° with respect to the static external magnetic field B_0 (Fig. 1.1a) leads to a time-averaging of the anisotropic interactions. E.g., CSA, the nuclear-electron interaction, is averaged to an isotropic value at sufficient spinning speed [30]. Except couplings that involve protons (e.g., $^1H-^1H$ or $^1H-^{13}C$), the effects of homonuclear and heteronuclear dipolar couplings (e.g., $^{13}C-^{13}C$ or $^{13}C-^{15}N$) are often removed under MAS [30], since the spinning speed can routinely be made greater than the magnitude of the coupling constant. Consequently, ssNMR spectra exhibit sharp line-widths similar to solution NMR (Fig. 1.1b). However, for slower spinning, spinning sidebands occur and are spaced at multiples of ω_r apart from the resonance line (Fig. 1.1b) [30]. Moreover, the common used spinning speeds (e.g., 11 kHz for a 4-mm rotor) are insufficient to average out the dipolar couplings that involve protons, i.e., $^1H-^1H$, $^1H-^{13}C$, and $^1H-^{15}N$ dipolar couplings.

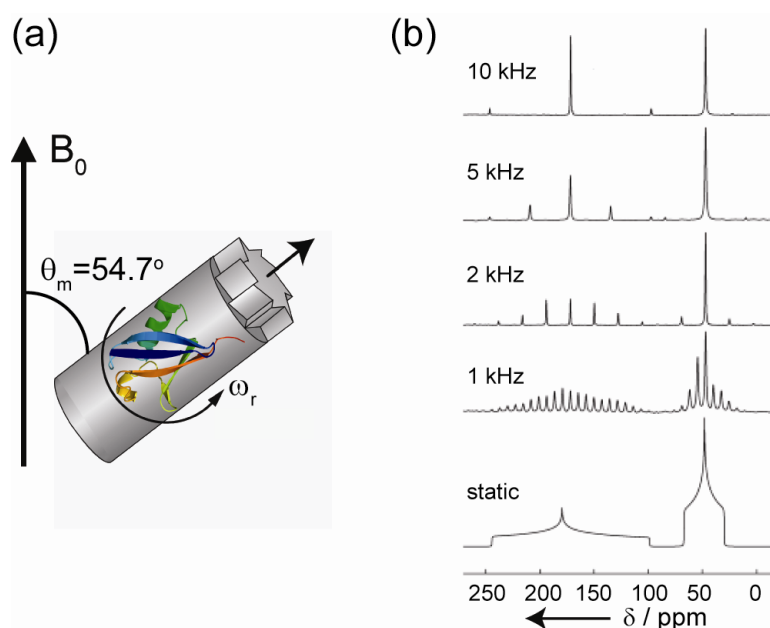


Figure 1.1. Magic-angle spinning. (a) The sample is rotating with a high frequency (ω_r) at an axis tilted by the magic angle (θ_m) of 54.7° with respect to the direction of the static external magnetic field B_0 . (b) ^{13}C ssNMR spectra of a uniformly ^{13}C -labeled (10%) glycine powder sample obtained with various MAS rates. Figure (b) obtained from Fig. 8 in Ref. [30].

To efficiently remove the heteronuclear ^1H - ^{13}C and ^1H - ^{15}N dipolar interactions, strong ^1H decoupling can be applied during evolution and detection periods for high-resolution ^{13}C or ^{15}N ssNMR spectra [35; 36]. However, as mentioned above, the strength of the dipolar coupling between two nuclei is proportional to the inverse cube of the internuclear distance (i.e., r^{-3}), which is the key for structural characterization including spectral assignments, distance restraints and torsion angle measurements. Therefore, to obtain structural information from MAS experiments, the homonuclear and/or heteronuclear dipolar couplings must be selectively restored. For this purpose, a number of recoupling pulse sequences exists, and can be used to reintroduce the interactions of interest [1; 37; 38].

For solution NMR on proteins, direct ^1H detection is widely used. However, ^1H spectra are normally of limited use in ssNMR spectroscopy due to the strong homonuclear dipolar interactions and small chemical shift dispersion of protein nuclei. Instead, direct detection on ^{13}C and ^{15}N are routinely employed in ssNMR spectroscopy. However, ^{13}C and ^{15}N are characterized by their significantly low natural abundances (1.11% for ^{13}C , 0.37% for ^{15}N) and low γ compared to those of ^1H . Regarding the low natural abundance of ^{13}C and ^{15}N nuclei in proteins, isotopic labeling is a prerequisite for

signal enhancement in ssNMR spectroscopy. For this purpose proteins are, for example, expressed in *Escherichia coli* (*E. coli*) cell cultures grown in a minimal medium containing uniformly [^{13}C]glucose ([U- ^{13}C]Glc) and ^{15}N - NH_4Cl as the sole carbon and nitrogen source, respectively. For sensitivity enhancement of nuclei with low γ (e.g., ^{13}C and ^{15}N), heteronuclear polarization transfer from the spin with high γ (e.g., ^1H) to the spin with low γ (e.g., ^{13}C or ^{15}N), i.e., cross-polarization (CP) [39], can be utilized. CP is an essential technique in ssNMR, and is established by applying radio-frequency (r.f.) pulses on both spin types, where the so-called Hartmann-Hahn condition needs to be fulfilled [40]. To reduce dependence on an exact field match, the strength of one of the two r.f. pulses may be ramped through the Hartmann-Hahn condition [41; 42]. In addition, CP is only effective for probing the rigid parts of one molecular system, because it relies on dipolar couplings. For highly mobile regions or regions exhibiting intermediate dynamics, complementary sequences, e.g., insensitive nuclei enhanced by polarization transfer (INEPT, J-coupling based transfer) [43; 44] are needed.

The first step towards structural investigation by ssNMR spectroscopy is to obtain the sequential resonance assignment of ^{13}C and ^{15}N resonances, i.e., correlating the observed spectral signals to individual nuclei in the primary sequence of the molecule under investigation. For this purpose, a set of 2D/3D ^{13}C - ^{13}C and ^{13}C - ^{15}N correlation ssNMR spectra has been developed based on uniformly or partially [^{13}C]glucose labeled peptides and proteins. Homonuclear ^{13}C - ^{13}C correlation experiments are often used for obtaining intra-residue and inter-residue correlations by longitudinal mixing sequences, e.g., proton-driven spin diffusion (PDS) [45; 46; 47] or dipolar-assisted rotational resonance (DARR) [48]. For example, intra-residue ^{13}C - ^{13}C correlations can be obtained in a 2D ^{13}C - ^{13}C PDS spectrum with a short mixing time (e.g., 20 ms on [U- ^{13}C]Glc-labeled samples), while inter-residue ^{13}C - ^{13}C correlations can be established in a 2D ^{13}C - ^{13}C PDS spectrum with a relatively long mixing time (e.g., 150 ms on [U- ^{13}C]Glc-labeled samples). Alternatively, heteronuclear ^{13}C - ^{15}N correlations (both intra-residue and inter-residue) can be obtained by selectively transferring magnetization from backbone ^{15}N to $^{13}\text{C}\alpha$ or $^{13}\text{C}'$ nuclei via SPECIFIC-CP [49; 50] through matching of Hartmann-Hahn condition [40]. SPECIFIC-CP is achieved by applying weak r.f. fields whose carrier frequency is centered on the interested resonance frequencies (~ 40 - 70 ppm for $^{13}\text{C}\alpha$ or ~ 165 - 185 ppm for $^{13}\text{C}'$) such that only $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i)$ or $^{15}\text{N}(i)$ - $^{13}\text{C}'(i-1)$ transfer is established. In addition to 2D NCA and NCO, 2D/3D NCACX and NCOCX experiments

are often used to obtain the sidechain resonances, established by one additional intra-residue ^{13}C - ^{13}C correlation step (via PDSO or DARR mixing). Moreover, through-space ^1H - ^1H contacts can be probed indirectly by detection on ^{15}N or ^{13}C nuclei directly bonded to ^1H using CHHC/NHHC schemes [51] with high spectral resolution. To ensure that the magnetization only transfers between ^1H and the directly bonded rare spins (i.e., ^{15}N or ^{13}C), short CP contact times need to be employed. Importantly in this thesis, under appropriate experimental conditions, by probing backbone ^1H - ^1H distances (via longitudinal ^1H - ^1H mixing) indirectly, NHHC spectra [51] carry information about the secondary structure of the molecule of interest. The magnetization transfer achieved using the above mentioned experiments is schematically represented in Fig. 1.2.

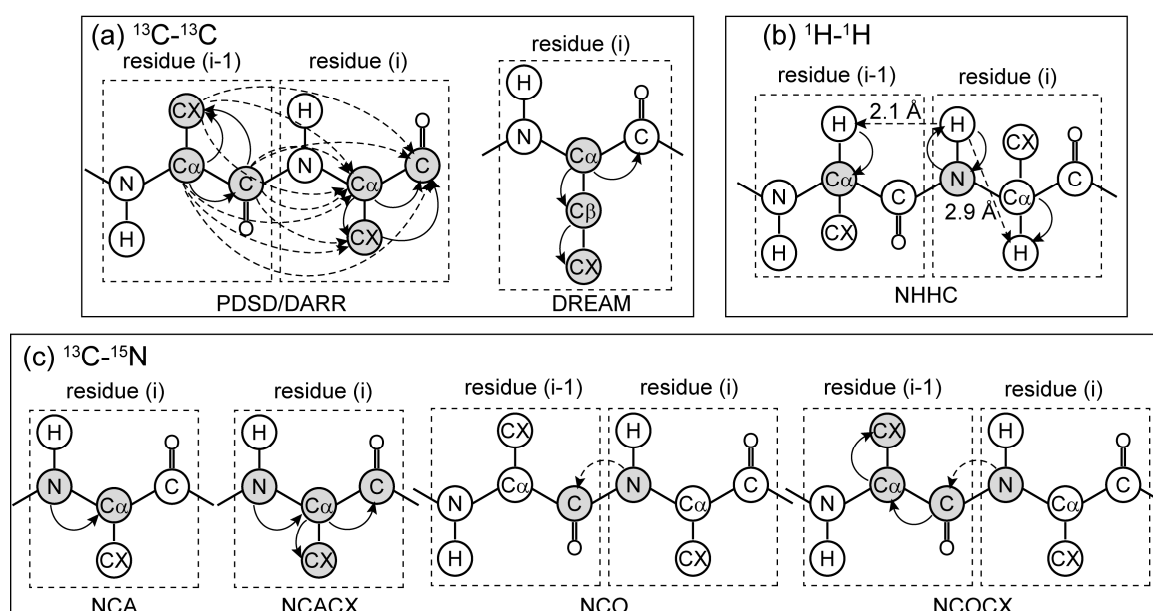


Figure 1.2. Schematic representation of magnetization transfer schemes achieved for probing (a) ^{13}C - ^{13}C , (b) ^1H - ^1H , and (c) ^{13}C - ^{15}N correlations. These experiments are commonly used in ssNMR spectroscopy, and aid in the assignment and identification of structural constraints. As one example for NHHC, the magnetization transfer between backbone $^1\text{HN}(i)$ and $^1\text{HC}\alpha(i)/(i-1)$ of two residues with a β -strand conformation (inter-residue $^1\text{HN}(i)$ - $^1\text{HC}\alpha(i-1)$ contacts are shorter than intra-residue $^1\text{HN}(i)$ - $^1\text{HC}\alpha(i)$ contacts) is shown in (b). The nuclei whose frequencies are detected are indicated by the gray circles, whereas the transfer of magnetization is indicated by the curved lines. Intra-residue and inter-residue magnetization transfer is denoted by solid and dashed arrows, respectively.

1.2 Sparse labeling schemes

In principle, ssNMR is a powerful means for structural investigation of biological macromolecules at atomic level. However, a major bottleneck in structural studies of proteins using ssNMR is the unambiguous and confident sequential resonance assignment which is often difficult due to poorly resolved and highly overlapped ssNMR spectra. Although the assignment problem for the entire protein sequence is largely facilitated by the [U- ^{13}C]Glc isotopic labeling approach, some spectroscopic problems are led. For example, the line broadening and spectral overlap due to abundant scalar J and dipolar couplings between adjacent ^{13}C nuclei, and make the unambiguous sequential assignment challenging. Additionally, the strong abundant dipolar couplings of adjacent ^{13}C nuclei in [U- ^{13}C]Glc-labeled samples make polarization transfer to remote spins difficult, a phenomenon called dipolar truncation [4; 52; 53]. The dipolar truncation effect lowers polarization transfer efficiency for long-range contacts, thus making the collection of long-range distance restraints challenging.

In order to simplify the spectrum and reduce the spectral overlap, amino acid specific or forward labeling is used, where the amino acids of interest are added to natural abundance bacterial growth medium. However, it is worth to note that, this approach is only specific for the amino acids at termini of biosynthetic pathways, otherwise, metabolic interconversion and thus scrambling can occur [54]. On the other hand, reverse labeling scheme can be employed, where certain amino acids are supplemented at natural abundance into an otherwise fully ^{13}C and ^{15}N labeled protein [55; 56]. The resonance assignment for few selective labels are thus obtained unambiguously. Such forward and reverse ^{13}C labeling schemes could also allow for the collection of correlations indicative of the global fold of a protein. Additionally, segmental labeling scheme [57] that isotopically labels a small segment of a large protein with ^{15}N and/or ^{13}C while leaving the rest of the protein unlabeled, is also a good choice to gain spectral simplification. However, the one-bond scalar J and dipolar couplings of adjacent ^{13}C nuclei are still present with the above mentioned labeling schemes.

Since the one-bond scalar J and dipolar couplings of adjacent ^{13}C nuclei are the main disadvantage, various advanced ^{13}C alternative labeling schemes can be employed. They are often termed as fractionally and extensively ^{13}C labeling schemes or sparse labeling schemes [58; 59; 60; 61], and the probability of adjacent ^{13}C atoms is minimized.

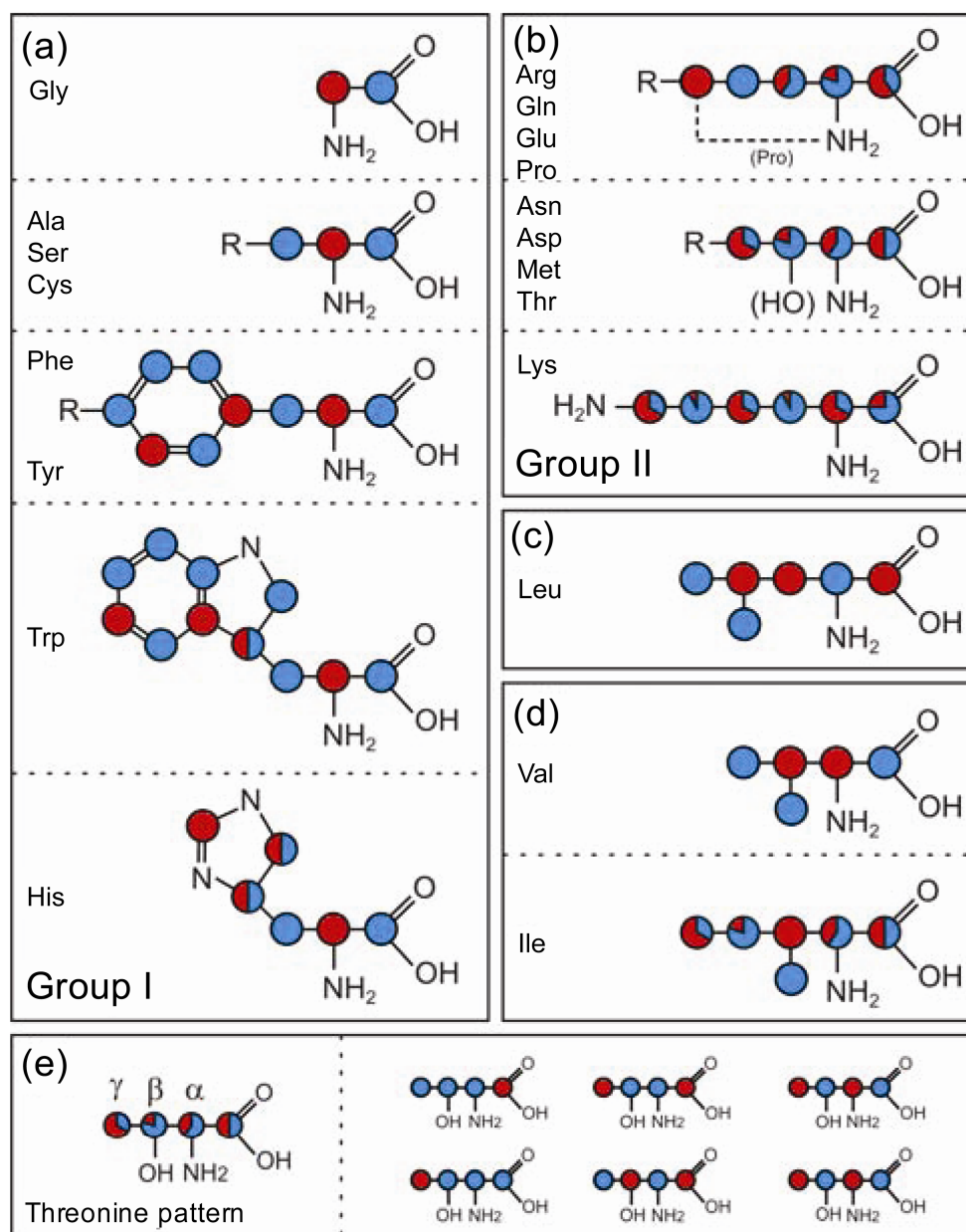


Figure 1.3. ^{13}C labeling patterns obtained by growth on media containing $[1,3-^{13}\text{C}]$ glycerol (in blue) and $[2-^{13}\text{C}]$ glycerol (in red). For residues with mixed labeling, the percentage of labeling is represented using relative blue/red coloring, except for the Trp C_γ and His C_δ and C_ϵ sites, for which the percentages could not be determined and a half/half distribution is displayed. The figure is adapted from Fig. 1 in Ref. [62].

Consequently, the spectral crowding is substantially alleviated, the sequential assignment and the collection of long-range distance restraints are facilitated.

One widely used sparse ^{13}C labeling scheme relies on the use of $[1,3-^{13}\text{C}]$ glycerol or $[2-^{13}\text{C}]$ glycerol as the carbon source for the biosynthetic production of amino acids [5; 59; 60]. The labeling schemes were first proposed to investigate protein dynamics via ^{13}C solution NMR relaxation analysis, and an *E. Coli* strain lacking succinate dehydrogenase

and malate dehydrogenase was used [59]. Subsequently, the labeling schemes of [1,3- ^{13}C]glycerol and [2- ^{13}C]glycerol using a normal *E. coli* strain were successfully applied for structural investigations of biomolecules by ssNMR [5; 6]. As shown in Fig. 1.3, the labeling patterns for the proteins produced from [1,3- ^{13}C]glycerol [6; 62] and [2- ^{13}C]glycerol [5; 6; 62], resemble an alternating “checkerboard” pattern. For amino acids produced from glycolytic three-carbon metabolites pyruvate (Ala and Val, Leu is the exception), 3-phosphoglycerate (3PG; Cys, Gly, Ser, and Trp), and phosphoenolpyruvate (PEP; Phe and Tyr), [1,3- ^{13}C]glycerol preferentially labels the methyl carbons [6; 62], whereas [2- ^{13}C]glycerol labels the backbone $\text{C}\alpha$ of these residues [5; 6; 62]. For amino acids (Fig. 1.3b,d) derived from the intermediates of the tricarboxylic acid cycle (TCA cycle), both the [1,3- ^{13}C]glycerol and [2- ^{13}C]glycerol labeling schemes lead to a mixed ^{13}C enrichment pattern, as illustrated with Thr (Fig. 1.3e) as an example [5; 6; 62].

This method is a very attractive isotope labeling approach for protein structural investigations by ssNMR. Due to the obtained alternating ^{13}C enrichment patterns [58; 59; 60; 61], the ssNMR spectral resolution is substantially enhanced and the resulting spectrum is simplified dramatically, which further facilitate the sequential assignment and allow for the observation of long-range distance restraints up to about 7 Å [6]. The high degree of unambiguity of long-range distance restraints allowed for the first structure determination of the α -spectrin SH3 domain by ssNMR [6].

Alternatively, two complementary sparse ^{13}C labeling schemes have been proposed for the use of dynamics investigation by solution NMR, i.e., [1- ^{13}C]glucose ([1- ^{13}C]Glc) [5; 61] and [2- ^{13}C]glucose ([2- ^{13}C]Glc) [61]. Similarly to the specifically ^{13}C labeled glycerol schemes (i.e., [1,3- ^{13}C]glycerol and [2- ^{13}C]glycerol), the proteins produced using [1- ^{13}C]Glc and [2- ^{13}C]Glc as the sole carbon source are ^{13}C labeled in an alternating ^{13}C enrichment pattern as well. Differently, the chemical symmetry of the glucose molecule (either [2- ^{13}C]Glc or [1- ^{13}C]Glc) reduces the maximum labeling level to 1/2, and the probability of being a ^{13}C spin pair is at most 1/4 even for two directly bonded labeled sites, which further alleviates the influence of the undesirable one-bond J couplings and dipolar couplings. In short, [1- ^{13}C]Glc [5; 61] and [2- ^{13}C]Glc [61] containing media produce proteins enriched at the methyl carbon positions and the backbone $\text{C}\alpha$ positions, respectively. For [2- ^{13}C]Glc-labeled proteins, $\text{C}\alpha$ is ^{13}C labeled without simultaneously ^{13}C labeling at $\text{C}\beta$ for 17 out of 20 residues, and no $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$ spin pairs is predicted [61]. For [1- ^{13}C]Glc-labeled proteins, all the amino acids

biosynthesized from glycolytic pathways are ^{13}C labeled at methyl carbons exclusively [5; 61]. For the purpose of studying protein dynamics using solution NMR, one of the labeling patterns where the C α s are expected to be ^{13}C labeled for proteins expressed in a $[2\text{-}^{13}\text{C}]\text{Glc}$ containing bacterial growth medium [61]. Likewise, one of the labeling patterns where the methyl carbons are predicted to be ^{13}C labeled for proteins produced using $[1\text{-}^{13}\text{C}]\text{Glc}$ were reported [61]. Moreover, the ^{13}C enrichment level at C α or methyl positions using either $[1\text{-}^{13}\text{C}]\text{Glc}$ or $[2\text{-}^{13}\text{C}]\text{Glc}$ as the carbon source for protein expression are experimentally measured using solution NMR [61]. In addition, for the purpose of structural studies of biomolecules by ssNMR, the distribution of ^{13}C labeled sites for proteins expressed using $[1\text{-}^{13}\text{C}]\text{Glc}$ were analyzed and verified by Hong [5]. However, it is worth to note that, alternative labeling pattern occurs for TCA cycle intermediates derived amino acids even after a single round of TCA cycle due to the symmetry of the TCA metabolite succinate, and particularly, more labeling patterns subsequently arise with different numbers of TCA cycles [61]. Nevertheless, a detailed and complete labeling pattern of the $[2\text{-}^{13}\text{C}]\text{Glc}$ and the $[1\text{-}^{13}\text{C}]\text{Glc}$ labeling schemes are still needed, which would be particularly useful especially for the use of ssNMR structural characterization of insoluble and non-crystalline biological assemblies. In the section 3.3 of Part I, an in-depth analysis of the $[1\text{-}^{13}\text{C}]\text{Glc}$ and $[2\text{-}^{13}\text{C}]\text{Glc}$ labeling patterns will be presented. Also the application to ssNMR will be discussed. The characteristic cross peaks expected and observed in 2D $^{13}\text{C}\text{-}^{13}\text{C}$ and $^{13}\text{C}\text{-}^{15}\text{N}$ ssNMR spectra will be exploited. Moreover, a sequential resonance assignment strategy based on 2D ssNMR spectra of $[1\text{-}^{13}\text{C}]\text{Glc}$ - and $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled samples will be proposed.

1.3 Parkinson's disease and α -synuclein fibrils

α -Synuclein (α S) is a 140 residue-long cytoplasmic protein that has been implicated in several neurodegenerative diseases, often termed as synucleinopathies, such as Parkinson's disease (PD), dementia with Lewy bodies, and multiple system atrophy [63; 64; 65]. PD, in particular, is the second most common neurodegenerative pathology after Alzheimer's disease, and neuropathologically characterized by the loss of dopaminergic neurons in the substantia nigra pars compacta of the brain and the formation of Lewy bodies (LBs) and Lewy neuritis (LNs) [66] whose major component is α S in the form of amyloid fibrils [67; 68; 69]. The aggregation of the 140 residue-long cytoplasmic protein α S is thus believed to play an important role in the etiology of PD.

The sequence of α S is divided into three domains [70]: a positively charged N-terminal region (residues 1-60), a central hydrophobic region known as NAC region (residues 61-95) that has been proposed to be important for aggregation, and an acidic C-terminal region (residues 96-140) enriched in Asp, Glu, and Pro (Fig. 1.4). α S has been classified as a "natively unfolded" or intrinsically disordered monomeric protein, which can adopt α -helical structure in solutions that contain lipid-mimetic detergent micelles or in the presence of lipid vesicles [70; 71; 72]. Upon aggregation, α S undergoes a conformational change into amyloid fibrils by intermediates of oligomers, ring-like oligomers, and protofibrils [73].

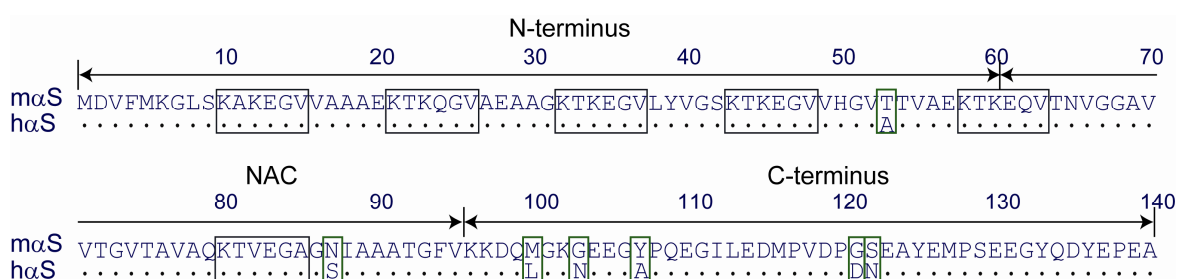


Figure 1.4. Primary sequence alignment of mouse α S (m α S) and human α S (h α S). Conserved residues are indicated by dots in the h α S sequence. The seven mutation sites are indicated with green boxes. The six highly repetitive, imperfect KTKEGV motifs are shown in black boxes. Residue numbers and N-terminal, NAC, and C-terminal regions are shown above the sequence.

Recently, an *in vivo* study indicated that endogenous h α S occurs physiologically as a folded tetramer with predominant helical secondary structure even without the

addition of lipid vesicles [74]. Around the same time, a heterologously expressed form of h α S with a 10-residue N-terminal extension was shown to form a stable tetramer in the absence of lipid bilayers or micelles [75]. Although there could be a big difference of the soluble form between *in vitro* and *in vivo* studies, *in vitro* fibrils have been shown to possess a morphology that closely resembles the one found in LBs of patients with PD [76; 77; 78].

Because LBs are a pathological hallmark of PD, extensive research has focused on the structural characterization of α S fibrils [44; 76; 77; 78; 79; 80; 81; 82; 83; 84]. Studies on h α S fibrils using ssNMR have led to substantial progress in the structural characterization at atomic level [44; 79; 80; 81; 82]. For example, two different morphologies of h α S fibrils, i.e., form A (twisted fibrils) and form B (straight fibrils), were identified and structurally investigated using ssNMR by Heise et al. [44]. 48 residues and 36 residues in the central region from Leu38 to Val95 were assigned and identified to be part of the central core for form A and form B of h α S fibrils, respectively. As shown in Fig. 1.5, for both forms of h α S fibrils studied by Heise et al. [44], the central core region consists mainly of well-defined β -strands.

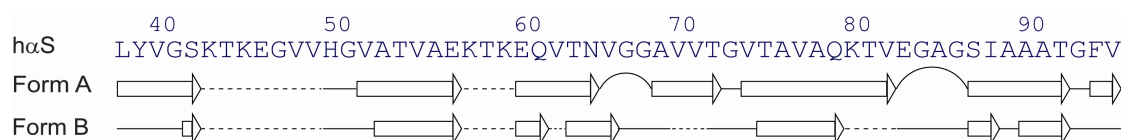


Figure 1.5. Secondary structure of the rigid core of h α S fibrils (form A and form B) studied by Heise et al. [44]. β -strands are indicated by white arrows, non- β -strand regions (kink, loop, or turn) are shown as a curve, non-assigned amino acids are marked as a dashed line, and straight lines are assigned amino acids for which a chemical shift analysis did not give well-defined torsion angles by TALOS. The figure is remade from Fig. 5 in Ref. [44].

Moreover, a possible structural model of h α S fibrils was proposed by Vilar et al. [80], based on the combination of hydrogen/deuterium (H/D) exchange NMR experiments, electron microscopy (EM) experiments and ssNMR data of a fragment of h α S fibrils (comprising residues 30-110). As shown in Fig. 1.6, the proposed structural model is a five-layered β -sandwich (β 1-loop- β 2-loop- β 3-loop- β 4-loop- β 5), generating five layers of parallel, in-register β -sheets. Two types of h α S fibrils, i.e., straight and twisted fibrils, were investigated and are illustrated in Fig. 1.6. In case of the straight fibril type, two protofilaments align with each other to form a fibril, which can align

again itself. For the twisted fibrils, two protofilaments twist around each other, and such a twisted filament twists again with another one. However, as commented by Comellas et al. [81], the fibril core of amyloid fibrils is generally well protected from H/D exchange, regardless of local secondary structure. In other words, less protected segments may be β -strand, but less stable and more exposed to solvent. Thus, H/D exchange cannot be used to draw definite structural conclusions.

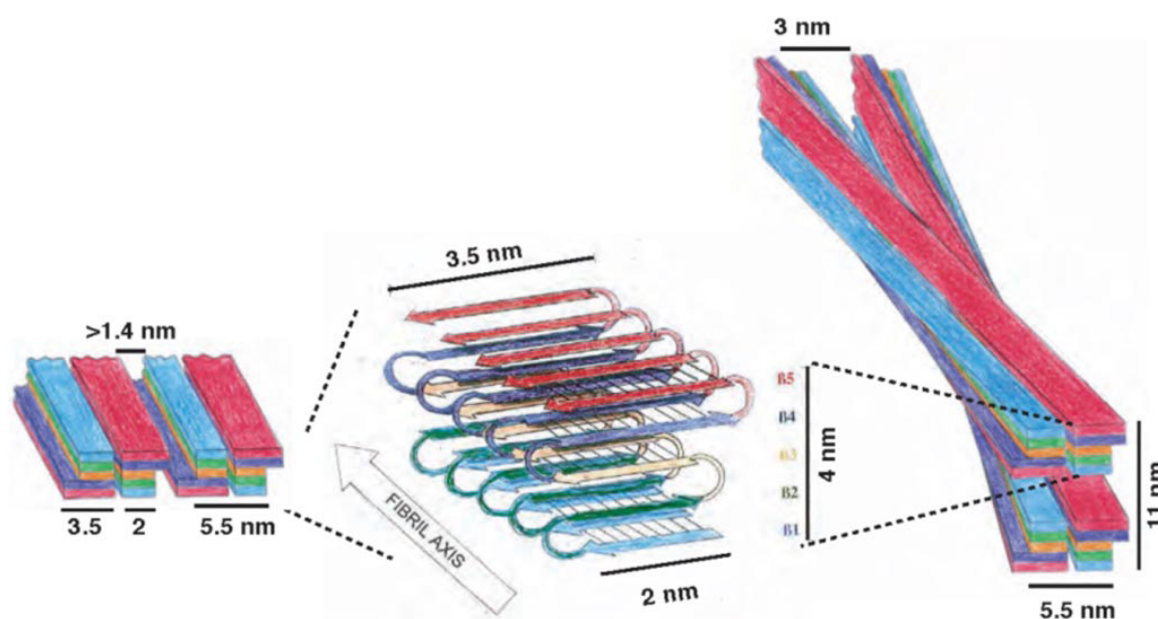


Figure 1.6. Proposed fold of h α S fibrils studied by Vilar et al. [80]. The proposed fold of a monomeric h α S molecule within a protofilament is shown in the center. The incorporation of a protofilament into the straight (left) and twisted (right) fibril type is indicated by a schematic drawing. The figure is adapted from Fig. 5 in Ref. [80].

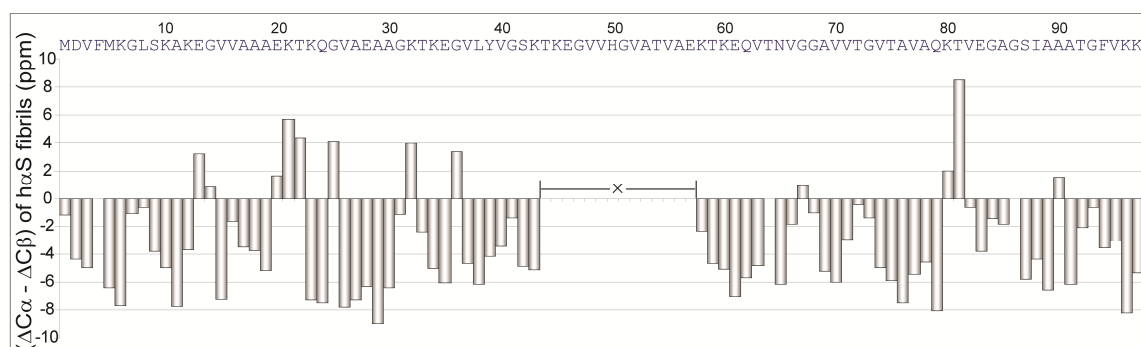


Figure 1.7. Secondary chemical shifts of the rigid core of the h α S fibrils studied by Gath et al. [82]. Non-assigned residues are marked with a cross, i.e., Thr44-Glu57. The figure is remade from Fig. 3 in Ref. [82].

More recently, Gath et al. [82] reported the sequential resonance assignment of the detected residues from Met1 to Lys97 for one polymorph of h α S fibrils. Residues from Met1 to Lys97 were identified to be mainly in β -sheet conformation (Fig. 1.7). For the first time, it was reported that the N-terminal residues are actually involved in the rigid part of h α S fibrils, which were postulated not to be part of the core in other studies [78; 80]. From the ssNMR spectra comparison, they assumed that the polymorph might be similar to the B-form described by Heise et al. [44].

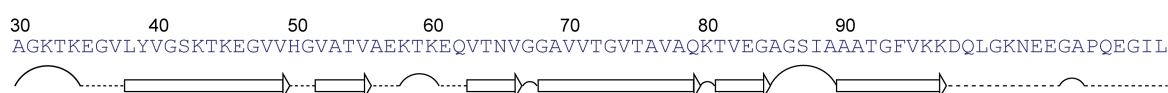


Figure 1.8. Representation of the secondary structure of the h α S fibrils studied by Comellas et al. [81] based on a TALOS+ analysis (arrows indicate β -strands; curved lines are turn or loop regions; dashed lines indicate no prediction). The figure is remade from Fig. 4d in Ref. [81].

Independently, the sequential resonance assignment of another morphology of h α S fibrils was achieved by Comellas et al. [81]. They obtained 91% of the ^{15}N and ^{13}C resonances of the detected residues between Leu38 and Lys96. As shown in Fig. 1.8, the secondary structure includes a repeated motif that is composed of a long β -strand followed by two short β -strands located from residues 38 to 66 and again from 68 to 96. The morphology studied by them is strongly different from those reported by Heise et al. [44] (i.e., form A and B) by a comparison of chemical shifts.

Table 1.1 α S Human-Mouse Variants. The table is adapted from Table. 1 in Ref. [85].

α S variants	name
Human	HHH
Human A53T	MHH
Human S87N	HMH
Human L100M-N103G-A107Y-D121G-N122S	HHM
Human A53T-S87N	MMH
Human A53T-L100M-N103G-A107Y-D121G-N122S	MHM
Human S87N-L100M-N103G-A107Y-D121G-N122S	HMM
Human A53T-S87N-L100M-N103G-A107Y-D121G-N122S (Mouse)	MMM

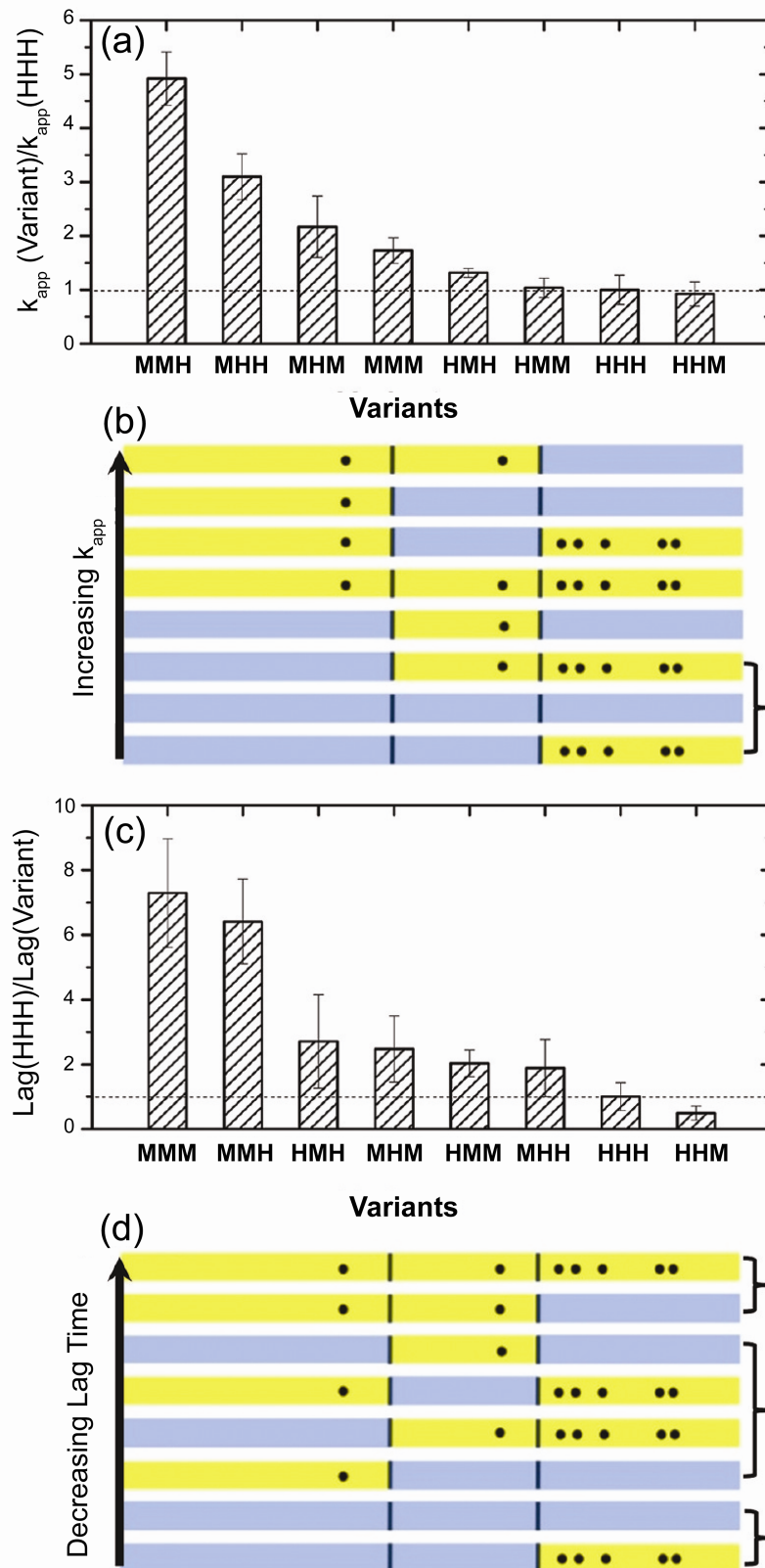


Figure 1.9. Fluorescence kinetic analysis of mouse-human chimera α S variants. (a) The ratio of apparent growth rates of α S variants over that of h α S (HHH). (b) Schematic representation of variant sequences in the order of the slowest to the fastest growth rates. The variants are color coded with the mouse region in yellow and the human region in blue. Dots represent substitutions from h α S to m α S. The bracket indicates the variants that have equivalent growth rates within the error bars. (c) The ratio of lag time of h α S (HHH) over that of α S variants. (d) Schematic representation of variant sequences in order of the longest to the

shortest lag time. Color code and dots are the same as in (b). The bracket indicates the variants that have equivalent lag times within the error bars. The figure is adapted from Fig. 3 in Ref. [85].

The primary sequence of m α S differs from h α S at seven positions (Fig. 1.4) [86]. An *in vitro* study showed that m α S has a “natively unfolded” structure in solution similar to h α S, whereas at elevated concentrations m α S forms amyloid fibrils with predominant β -sheet secondary structure much more rapidly than its human counterpart [87]. The fibrillization occurs via a nucleation-dependent mechanism, which is typically characterized by the presence of a lag phase followed by an exponential growth phase. The *in vitro* study showed that the lag phase for m α S was shorter than that for h α S [87]. Moreover, in order to provide the critical residues or regions that lead to faster nucleation and elongation of m α S compared to that of h α S, a detailed sequence-dependent fibrillization difference between m α S and h α S was recently investigated by Kang et al. [85]. Systematically, they designed seven variants of h α S and m α S, either by a mutation at the single residue level (i.e., Ala53Thr and Ser87Asn), or by a five-fold mutation (Leu100Met, Asn103Gly, Ala107Tyr, Asp121Gly, and Asn122Ser) (Table 1.1). The variants contain either one mutation of the human to mouse region (e.g., MHH, HMH, or HHM) or two mutations (HMM, MHM, or MMH) or all three mutations (MMM) (Table 1.1). As revealed from Fig. 1.9b, it is clear that the Ala53Thr substitution dominates the aggregation growth rates, and the data presented in Fig. 1.9d suggested that the combination of the Ala53Thr and Ser87Asn mutations plays a significant role in determining the length of the lag phase [85].

An understanding of the different aggregation kinetics between m α S and h α S is desirable, since it was reported that neither PD symptoms nor LBs have been observed in aged mice [87]. Interestingly, it was recently reported that in wild-type nontransgenic mice, a single intrastriatal inoculation of recombinant m α S preformed fibrils initiates a neurodegenerative cascade characterized by the accumulation of intracellular LBs/LNs pathology, selective loss of dopamine neurons in the substantia nigra pars compacta and motor deficits [88].

It should be worthwhile to investigate the structure of m α S fibrils at atomic level and then compare to the structure of h α S fibrils in order to gain further insights into the largely different aggregation kinetics. However, much less is known about the structural properties of m α S fibrils compared to h α S fibrils. Rochet et al. has demonstrated that

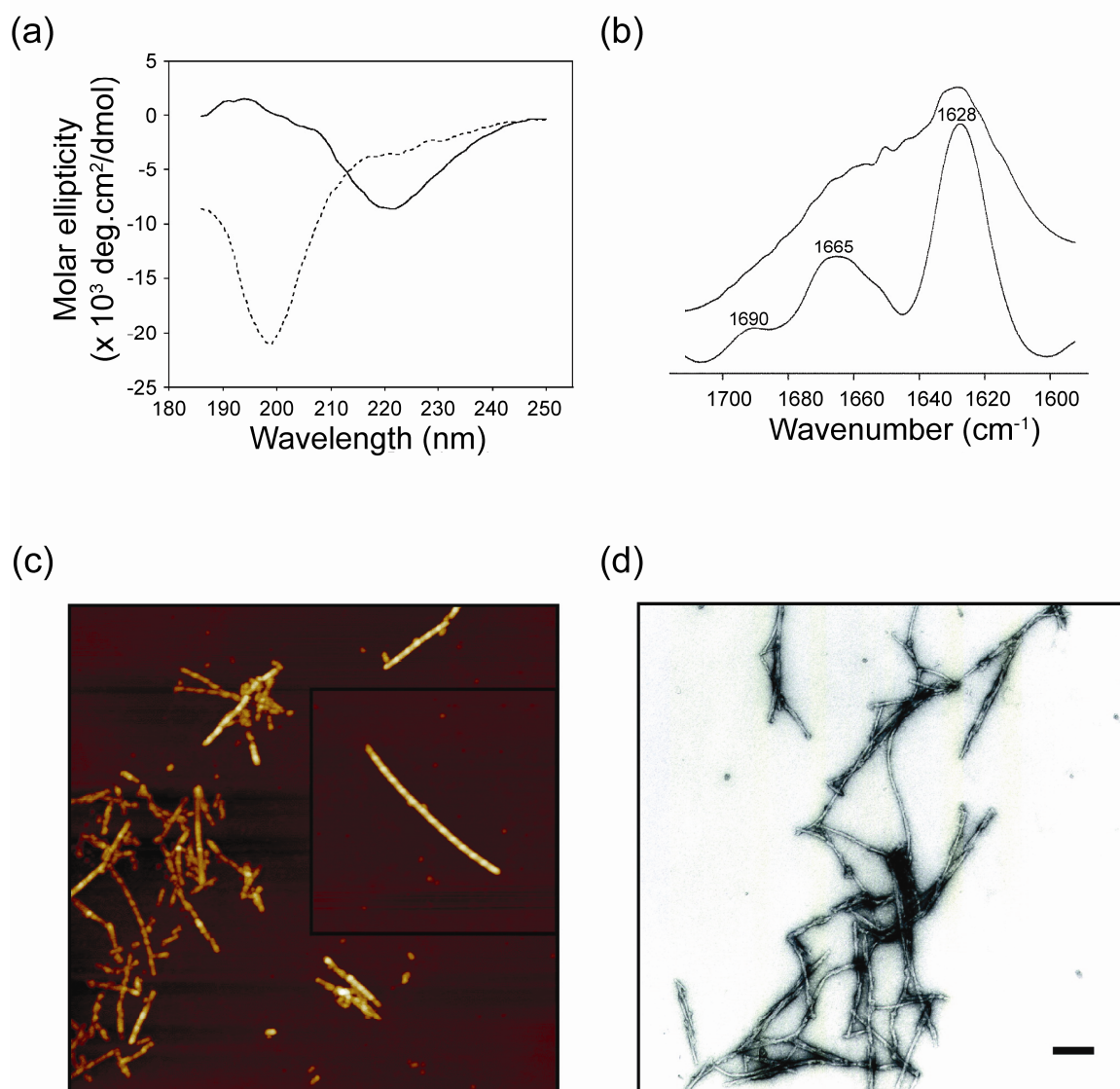


Figure 1.10. Fibrils of $m\alpha S$ have the properties of amyloid [87]. (a) Far-UV CD spectra of the supernatant (dashed line) and resuspended pellet (solid line) obtained by centrifugation of an incubated sample of $m\alpha S$. (b) Analysis of fibrillar $m\alpha S$ by FTIR spectroscopy. The undeconvoluted and deconvoluted spectra appear at the top and bottom, respectively. (c) AFM images showing fibrils of $m\alpha S$. Main image, $2 \mu\text{m} \times 2 \mu\text{m}$; inset image, $1 \mu\text{m} \times 1 \mu\text{m}$. (d) EM image of fibrillar $m\alpha S$. Scale bar, 200 nm. The figure is adapted from Fig. 2 in Ref. [87].

$m\alpha S$ fibrils are also characterized by a predominance of β -sheet secondary structure by means of Far-UV Circular Dichroism spectroscopy (Far-UV CD, Fig. 1.10a) and Fourier transform infrared spectroscopy (FTIR, Fig. 1.10b) [87]. Moreover, the dimensions and morphology of $m\alpha S$ fibrils were similar to those reported for the fibrils of $h\alpha S$, Ala53Thr $h\alpha S$, and Ala30Pro $h\alpha S$. The fibrils were straight, rigid, and unbranched, with lengths ranging from 0.1 to $3 \mu\text{m}$ as observed using atomic force microscopy (AFM) and EM (Fig. 1.10c-d), while the mean fibril diameter was $12.1 \pm 0.9 \text{ nm}$ as determined by EM

(Fig. 1.10d) [87]. However, the extent of the core region and site-specific structural information remained absent. The fact that, so far no structural details of m α S fibrils and their differences to h α S fibrils were available motivated the present study.

In the Part I of this thesis, the structural characterization of m α S fibrils was studied using state-of-the-art ssNMR. As a major bottleneck in structural studies of fibrillar proteins by ssNMR remains the unambiguous and confident sequential resonance assignment which is often difficult due to poorly resolved and highly crowded ssNMR spectra. Furthermore, structural polymorphism can lead to the phenomenon that different sample preparations exhibit slightly different ssNMR spectra. The combination of a highly reproducible sample preparation, tailored isotope-labeling schemes, and state-of-the-art ssNMR methods enabled us to obtain the near-complete sequential resonance assignment and a subsequent structural characterization of the fibril core of m α S fibrils. For this purpose we prepared [U- ^{13}C]Glc-labeled m α S fibrils, and two complementary sparse ^{13}C enriched variants of m α S fibrils. Bacterial growth in medium containing [1- ^{13}C]Glc or [2- ^{13}C]Glc results in highly diluted ^{13}C labeling in proteins with only one out of six carbons labeled [58; 61; 89]. This leads to enhanced resolution and a reduced number of cross peaks which facilitates the sequential resonance assignment considerably [89]. As an illustration, the near-complete *de novo* sequential assignment of m α S was obtained, 96% of backbone amide ^{15}N and 93% of all ^{13}C atoms of the detected residues from Gly41 to Val95. Moreover, six β -strands were identified to be within the fibril core of m α S fibrils by secondary chemical shift and NHHc [51] analysis. Additionally, by measurement of intermolecular ^1H - ^1H correlations using the NHHc scheme on a mixed 1:1 ^{13}C : ^{15}N ([M- $^{13}\text{C}/^{15}\text{N}$]-) labeled variant [90] of m α S fibrils we could deduce the supra-molecular arrangement of m α S fibrils. We found that they are stacked parallel, and in-register. This result is in agreement with our recent investigation of h α S fibrils prepared under identical fibrillization conditions to the m α S fibrils studied here and which were also found to adopt a parallel, in-register β -sheet arrangement [91].

2 Materials and methods

2.1 Sample preparation and morphological characterization

2.1.1 α -Synuclein protein expression in *E. coli* and purification

For the sequential resonance assignment, three samples were used in this work, i.e., [U-¹³C]Glc-, [1-¹³C]Glc-, and [2-¹³C]Glc-labeled α S samples (produced by Dr. Stefan Becker and Karin Giller at the Max Planck Institute for Biophysical Chemistry). These three samples were uniformly ¹⁵N-labeled. The α S gene was expressed in *E. coli* using plasmid pT7-7 encoding for the protein (courtesy of the Lansbury laboratory, Harvard Medical School, Cambridge, MA). The protein was purified according to established protocols [92]. Following transformation, BL21 (DE3) cells were grown overnight in a minimal medium in the presence of ampicillin (100 μ g/ml, supplemented with ¹⁵N-NH₄Cl as the sole nitrogen source, and one of the following three compounds: [U-¹³C]Glc, [1-¹³C]Glc, and [2-¹³C]Glc as the sole carbon source. Cells were induced with IPTG, cultured at 37 °C for four hours and harvested by centrifugation at 6000 rpm (7459 \times g) for 10 minutes. The cell pellet was resuspended in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA and 1 mM PMSF, and lysed by multiple freeze–thaw cycles and sonication. The cell suspension was boiled for 20 minutes and centrifuged at 20,000 rpm (48384 \times g) for 30 minutes at 4 °C. Streptomycin sulfate was added to the supernatant from the previous step to a final concentration of 10 mg/ml and the mixture was stirred for 15 minutes at 4 °C. After centrifugation at 20,000 rpm (48384 \times g), ammonium sulfate (0.36 g/ml) was added to the collected supernatant. The solution was stirred at 4 °C for 30 minutes and centrifuged again at 20,000 rpm (48384 \times g) for 30 minutes at 4 °C. The protein pellet was resuspended in 25 mM Tris–HCl (pH 7.7), and loaded onto an HQ/M-column on an Applied Biosystems BIOCAD (Applied Biosystems, Foster City, CA) workstation. α S was eluted at ~300 mM NaCl with a salt gradient from 0 mM to 600 mM NaCl. The pure α S was dialyzed overnight against the desired buffer. Protein purity was judged by PAGE, ESI-MS, and analytical gel-filtration, and the concentration was estimated from the absorbance at 275 nm using an extinction coefficient of 5600 M⁻¹ cm⁻¹.

[M-¹³C/¹⁵N]-labeled mαS fibrils were prepared by aggregating an equimolar mixture of uniformly ¹³C-labeled and ¹⁵N-labeled mαS. The uniformly ¹³C-labeled and ¹⁵N-labeled mαS were produced with the same protocol as mentioned above by supplementing with [U-¹³C]Glc as the sole carbon source, and ¹⁵N-NH₄Cl as the sole nitrogen source, respectively.

2.1.2 Fibrillization *in vitro*

Fibrillization of mαS was achieved according to the protocol previously developed for hαS [44]. In short, mαS solutions were incubated with protein concentrations of 200 μM in 25 mM Tris–HCl buffer, pH 7.5 containing 0.02% NaN₃ at 37 °C with continuous stirring with micro stir bars at 300 rpm until the concentration of fibrillized protein reached a steady state according to a thioflavin T (ThT) fluorescence assay [93]. Briefly, 5-10 μl aliquots were withdrawn from the incubations and diluted into 2 ml of 50 μM ThT in 50 mM Glycine buffer, pH 8.0. ThT fluorescence was measured in a Varian Cary Eclipse spectrofluorimeter, with an excitation wavelength of 446 nm. Fluorescence emission from 460 to 600 nm was collected. Fluorescence at 480 nm was employed for determination of the relative content of mαS fibrils in the sample.

2.1.3 Electron microscopy

The morphology of the mαS fibrils was monitored by negative stained EM at room temperature (performed by Dr. Dietmar Riedel and Gudrun Heim at the Max Planck Institute for Biophysical Chemistry). MαS fibrils, resuspended from the pellet, were diluted by Tris-HCl buffer (25 mM, pH 7.5), to a concentration of 25 μM and then deposited on carbon-coated 200 mesh copper grids (Electron Microscopy Sciences). Grids were stained with 1% (w/v) uranyl acetate aqueous (Electron Microscopy Sciences). Specimens were evaluated using a FEI CM120 transmission electron microscope, operated at 120 kV. Pictures were taken using a TemCam 224A slow scan CCD camera (TVIPS, Gauting, Germany) at a defocus of -3μm.

2.2 Solid-state NMR experiments and analysis

All ssNMR experiments were conducted using 4-mm and 3.2-mm triple resonance (^1H , ^{13}C , ^{15}N) probe heads at static magnetic fields of 20.0 Tesla, 18.8 Tesla, and 14.1 Tesla, corresponding to 850 MHz, 800 MHz, and 600 MHz ^1H resonance frequencies (Bruker Biospin, Germany). The chemical shifts of ^{13}C and ^{15}N were calibrated either with DSS as an internal reference [94] or with adamantane as an external reference [95]. Sample temperatures were determined by the position of the water ^1H resonance using the relation $\delta(\text{H}_2\text{O}) = 7.83 - T/96.9$ [31]. All experiments were carried out at a sample temperature of around +278 K, and MAS rates between 11 and 18 kHz. An initial ramped CP [41; 42] was used to transfer the magnetization from ^1H to ^{13}C or ^{15}N with contact times between 400-1200 μs , and ^{15}N to ^{13}C transfer was achieved using SPECIFIC-CP [49] with contact times of 3-6 ms. High-power proton decoupling using SPINAL-64 [36], or CW [96] was applied during evolution and detection periods with r.f. amplitudes of 69-83 kHz.

2.2.1 Experiments on uniformly [^{13}C]glucose labeled mouse α -synuclein fibrils

2.2.1.1 2D ^{13}C - ^{13}C and ^{15}N - ^{15}N correlation experiments

For the sequential assignment, 2D ^{13}C - ^{13}C correlation experiments were conducted using PDSM with mixing times of 20, 80, and 150 ms for detecting intra-residue and inter-residue ^{13}C - ^{13}C correlations (pulse sequence: Fig. A1). In order to detect one-bond correlations of rigid protein segments, a dipolar recoupling enhancement through amplitude modulation (DREAM) [97] spectrum was recorded with a recoupling period of 4 ms at a spinning speed of 18 KHz in the 3.2-mm rotor (pulse sequence: Fig. A2). Moreover, double-quantum single-quantum correlation ((2Q,1Q) correlation) spectrum was recorded with the supercycled POST-C5 (SPC5) scheme [98], to separate the signals overlapped in diagonal region (pulse sequence: Fig. A3). Additionally, in order to obtain sequential connections of amide ^{15}N resonances, 2D ^{15}N - ^{15}N correlation experiments were conducted via PDSM [99] with mixing times of 6 s and 10 s (pulse sequence: Fig. A1).

2.2.1.2 2D/3D ^{15}N - ^{13}C correlation experiments

For intra-residue correlations, 2D NCA and NCACX spectra were recorded. The ^{13}C - ^{13}C polarization transfer of the 2D NCACX spectra was established via PDSM with a mixing time of 60 ms or via DARR [48] with a mixing time of 50 ms. For inter-residue correlations, 2D NCO and NCOCX spectra were measured. The ^{13}C - ^{13}C polarization transfer of the 2D NCOCX spectra was established via PDSM with mixing times of 25 and 50 ms, and via DARR [48] with a mixing time of 50 ms. The cross peak patterns observed via PDSM and DARR are identical, slightly shorter mixing times were employed for DARR than for PDSM, due to the more efficient magnetization transfer of DARR. In order to reduce spectral overlap that appear in 2D ssNMR spectra, 3D NCACX and NCOCX spectra were conducted for intra-residue and inter-residue correlations, respectively. The ^{13}C - ^{13}C polarization transfer was established via PDSM with a mixing time of 50 ms for both spectra. The pulse sequences for 2D NCA/NCO and 2D NCACX/NCOCX spectra (via PDSM mixing) are depicted in appendix Fig. A4 and Fig. A5, respectively.

2.2.1.3 Indirect detection of ^1H - ^1H correlations

In order to characterize the secondary structure of m α S fibrils, backbone ^1H - ^1H distances were probed indirectly by using the NHHC scheme [51]. The pulse sequence is depicted in appendix Fig. A6. The contact time of the initial CP from ^1H to ^{15}N was set to 600 μs to allow that polarization transfers from all ^1H to amide ^{15}N . A longitudinal ^1H - ^1H mixing time (t_{HH}) of 150 μs was used to detect the nearest ^1H - ^1H correlations. To ensure that polarization transfer only occurs within bonded ^1H - ^{13}C or ^1H - ^{15}N pairs, short contact times were set to for the CP from ^{15}N to ^1H and from to ^1H to ^{13}C , i.e., $t_{\text{HN}} = 200 \mu\text{s}$ and $t_{\text{HC}} = 200 \mu\text{s}$.

2.2.2 Experiments on [2- ^{13}C]glucose labeled mouse α -synuclein fibrils

For 2D ^{13}C - ^{13}C correlation experiments, PDSM were conducted with different mixing times for detecting intra-residue and inter-residue ^{13}C - ^{13}C correlations. 2D ^{15}N - ^{15}N correlation experiments were conducted via PDSM with a mixing time of 6 s to obtain sequential connections of amide ^{15}N resonances. For ^{13}C - ^{15}N correlation experiments, 2D

spectra of NCA and NCO were recorded to detect intra-residue correlations (i.e., $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i)$), and inter-residue correlations (i.e., $^{15}\text{N}(i)$ - $^{13}\text{C}'(i-1)$), respectively. Similarly to $[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils, backbone ^1H - ^1H distances were probed indirectly by using the NHHC scheme [51] to characterize the secondary structure of $\text{m}\alpha\text{S}$ fibrils. The contact time of the initial CP was set to 400 μs . A longitudinal ^1H - ^1H mixing time of 175 μs was used. Short contact times for the second and third CP (i.e., from ^{15}N to ^1H , and from ^1H to ^{13}C), $t_{\text{HN}} = 400 \mu\text{s}$ and $t_{\text{HC}} = 200 \mu\text{s}$, were set.

2.2.3 Experiments on $[\text{1-}^{13}\text{C}]\text{glucose}$ labeled mouse α -synuclein fibrils

Likewise, 2D ^{13}C - ^{13}C PDSO correlation experiments were conducted with different mixing times for detecting intra-residue and inter-residue ^{13}C - ^{13}C correlations. For ^{13}C - ^{15}N correlation experiments, 2D spectra of NCA and NCO were recorded to detect intra-residue correlations (i.e., $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i)$), and inter-residue correlations (i.e., $^{15}\text{N}(i)$ - $^{13}\text{C}'(i-1)$), respectively.

2.2.4 Experiment on mixed 1:1 ^{13}C : ^{15}N labeled mouse α -synuclein fibrils

To characterize the supra-molecular structure of $\text{m}\alpha\text{S}$ fibrils, an NHHC spectrum [51] was recorded on mixed 1:1 ^{13}C : ^{15}N labeled ($[\text{M-}^{13}\text{C}/^{15}\text{N}]$ -labeled) $\text{m}\alpha\text{S}$ fibrils. A longitudinal ^1H - ^1H mixing time of 500 μs was used to detect cross peaks between ^{13}C and ^{15}N that correspond to intermolecular ^1H - ^1H distances up to about 3.5 \AA . The contact time of the initial CP from ^1H to ^{15}N was set to 700 μs to allow that polarization transfers from all ^1H to amide ^{15}N . To ensure that polarization transfer only occurs within bonded ^1H - ^{15}N or ^1H - ^{13}C pairs, short contact times for the CP from ^{15}N to ^1H and from ^1H to ^{13}C , $t_{\text{HN}} = 200 \mu\text{s}$ and $t_{\text{HC}} = 200 \mu\text{s}$ were used.

All the experiments used for sequential resonance assignments and supra-molecular structural characterization of $\text{m}\alpha\text{S}$ fibrils together with their detailed acquisition parameters are summarized in Table A1 (appendix A). Data were processed with Topspin (Bruker Biospin, Germany). All the spectra were zero-filled, apodized with squared sine bell window functions along direct and indirect dimensions, and then Fourier

transformed. Specific processing parameters are indicated in individual figure legends. The assignment and linewidths were analyzed using SPARKY version 3.1 (T. D. Goddard & D. G. Kneller, University of California).

3 Results and discussion

3.1 Morphological characterization by electron microscopy and initial comparison to human α -synuclein fibrils

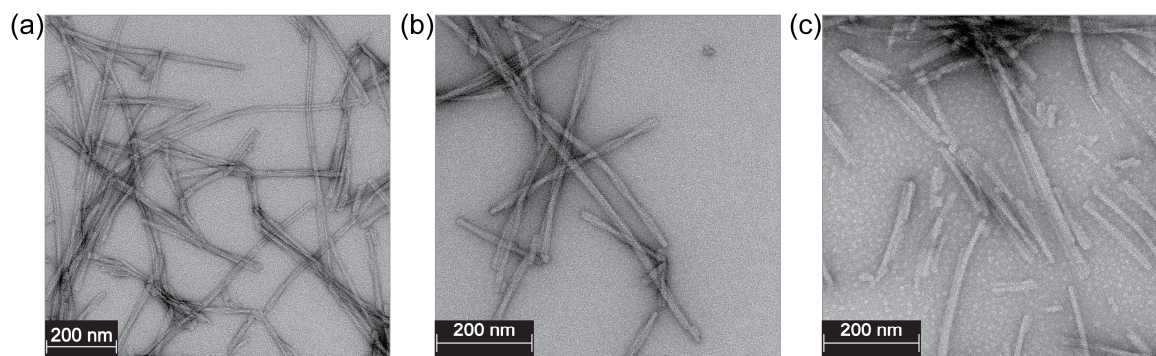


Figure 3.1. EM micrographs of (a) [U- ^{13}C]Glc-, (b) [1- ^{13}C]Glc-, and (c) [2- ^{13}C]Glc-labeled $m\alpha\text{S}$ fibrils used for ssNMR experiments.

In this study, [U- ^{13}C]Glc-, [1- ^{13}C]Glc-, and [2- ^{13}C]Glc-labeled $m\alpha\text{S}$ fibrils were prepared under identical conditions to obtain the sequential resonance assignment. The morphology of the three samples was then monitored by EM. The EM micrographs of [U- ^{13}C]Glc-, [1- ^{13}C]Glc-, and [2- ^{13}C]Glc-labeled $m\alpha\text{S}$ fibrils are shown in Fig. 3.1a, b, and c, respectively. The $m\alpha\text{S}$ fibrils were all well ordered and long straight fibrils with a diameter around 160 Å dominate all three samples. Rarely, twisted fibrils were observed in [U- ^{13}C]Glc- (Fig. 3.1a), [1- ^{13}C]Glc- (Fig. 3.1b) or [2- ^{13}C]Glc-labeled $m\alpha\text{S}$ fibrils (Fig. 3.1c). The highly similar morphology of the three samples indicates high reproducibility of our sample preparation which is an essential requirement if different labeling schemes are to be used for ssNMR. However, both straight and helically twisted fibrils of $m\alpha\text{S}$ were observed by EM in the recent study by Kang et al. [85]. The observed differences of the $m\alpha\text{S}$ fibril morphology might be due to different fibrillization conditions, and suggests the existence of two types of $m\alpha\text{S}$ fibrils (straight and twisted). In the case of $h\alpha\text{S}$ fibrils, the overall morphology was demonstrated to be strongly dependent on the fibrillization conditions, and both straight and twisted fibrils were observed and characterized [44]. Nevertheless, the highly uniform morphology of $m\alpha\text{S}$ fibrils obtained

by our preparation indicates high molecular homogeneity of our samples, and indeed no detectable polymorphism was observed in our recorded ssNMR spectra (see below). The absence of polymorphism allowed us to obtain sequential resonance assignment of the core region of $m\alpha S$ fibrils with straight morphology.

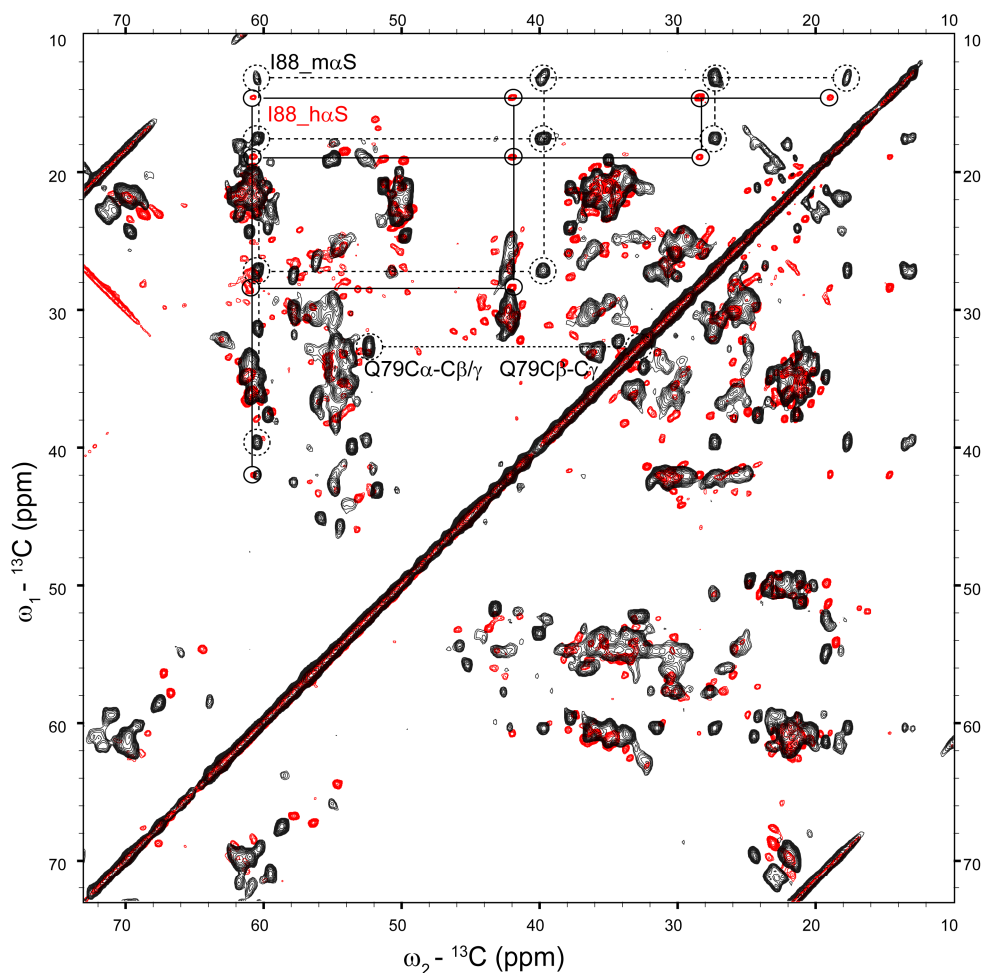


Figure 3.2. Initial comparison of 2D ^{13}C - ^{13}C PDSD spectra of $[\text{U}-^{13}\text{C}]$ Glc-labeled $m\alpha S$ fibrils (black) and $[\text{U}-^{13}\text{C}]$ Glc-labeled $h\alpha S$ fibrils (red, reproduced from Ref. [91]). Correlations of residue Ile88 for $m\alpha S$ and $h\alpha S$ are illustrated with dashed and solid lines, respectively. $m\alpha S$ and $h\alpha S$ fibrils were obtained under identical fibrillization conditions. The black spectrum was recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. The black spectrum was zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

As mentioned above, the primary sequence of $m\alpha S$ differs from $h\alpha S$ at seven positions (Fig. 1.4) [86]. Despite the highly conserved sequence, the aggregation kinetics of $m\alpha S$ and $h\alpha S$ are remarkably different [85; 87]. $m\alpha S$ aggregates about twice faster and the lag phase is reduced by almost one order of magnitude [85]. To characterize the structural differences between $m\alpha S$ and $h\alpha S$ fibrils at the atomic level, a set of high

resolution ssNMR spectra was recorded on m α S fibrils and compared to existing data for h α S fibrils that were prepared under identical conditions [91]. As an example, Fig. 3.2 shows the comparison of 2D ^{13}C - ^{13}C PDSO spectra with a mixing time of 20 ms recorded on h α S (red) and m α S fibrils (black). The h α S fibrils considered here exhibit the same morphology as m α S fibrils (i.e., straight). From the spectral comparison, a difference between m α S fibrils and h α S fibrils is obvious, as illustrated with the straightforward Ile88 assignment in Fig. 3.2. Since the amino acid sequence of m α S differs from h α S at position 87 (Ser vs. Asn) (Fig. 1.4), different chemical shifts of Ile88 due to the disturbance of the chemical environment by the Ser87Asn mutation are expected. A more detailed comparison between m α S and h α S fibrils requires the sequential resonance assignment of m α S fibrils as described in the following sections of Part I.

3.2 Sequential resonance assignment using uniformly [^{13}C]glucose labeled mouse α -synuclein fibrils

Overall, the procedure to obtain sequential resonance assignments is to identify spin systems, assign the spin systems to an amino acid type and link the spin systems according to the primary sequence of the protein. Here, a sequential resonance assignment strategy using 2D ^{13}C - ^{15}N and ^{13}C - ^{13}C correlation spectra recorded on [^{13}C]Glc-labeled m α S fibrils is illustrated in Fig. 3.3. The recorded 2D NCACX, NCOCX, and 2D ^{13}C - ^{13}C PDSO spectra were well resolved, and the ^{13}C and ^{15}N linewidths (full width at half height) were found to be ~ 0.5 ppm and ~ 1.3 ppm, respectively. The 2D ^{13}C - ^{13}C PDSO spectrum with a short mixing time of 20 ms (red spectrum in Fig. 3.3) is typically dominated by intra-residue ^{13}C - ^{13}C correlations, which helped us to identify spin systems. Inter-residue ^{13}C - ^{13}C cross peaks, on the other hand, were readily observed in the 2D ^{13}C - ^{13}C PDSO spectrum with a long mixing time of 150 ms (green spectrum in Fig. 3.3) which enabled us to obtain sequence-specific resonance assignment. Similarly, the 2D NCACX spectrum (purple spectrum in Fig. 3.3) with a PDSO mixing time of 60 ms establishes intra-residue ^{15}N - ^{13}C correlations, i.e., $^{15}\text{N}(i)$ - $^{13}\text{CX}(i)$ (where CX stands for any carbon, e.g., C', C α , C β , C γ , etc.). In contrast, the 2D NCOCX spectrum with a PDSO mixing time of 50 ms (blue spectrum in Fig. 3.3) generally yields inter-residue correlations between nitrogen atoms and carbon atoms of the preceding residue, i.e., $^{15}\text{N}(i)$ - $^{13}\text{CX}(i-1)$.

As an example, the sequential resonance assignment of Gly41-Ser42-Lys43 is illustrated in Fig. 3.3. In the 2D ^{13}C - ^{13}C PDSO spectrum with a mixing time of 20 ms, the intra-residue correlations at 58.7-67.7, 58.7-171.5, and 67.7-171.5 ppm could be easily identified to be Ser spin system, according to its characteristic chemical shifts. Subsequently, sequential resonance assignments involving $^{13}\text{C}\alpha(i)$ - $^{13}\text{C}\alpha(i\pm 1)$ correlation (at 58.7-47.9) and $^{13}\text{C}\beta(i)$ - $^{13}\text{C}\alpha(i\pm 1)$ correlation (at 67.7-47.9 ppm) were then observed in the 2D ^{13}C - ^{13}C PDSO spectrum with a mixing time of 150 ms, and were unambiguously identified to be the sequential correlation of Gly41-Ser42, according to the primary sequence of m α S. Moreover, the correlations of $^{15}\text{N}(i)$ - $^{13}\text{CX}(i)$ and $^{15}\text{N}(i)$ - $^{13}\text{CX}(i-1)$ can be easily connected in the spectra of NCACX and NCOCX from the ^{13}C - ^{13}C correlation spectra, and allowed us to obtain more sequential resonance assignments.

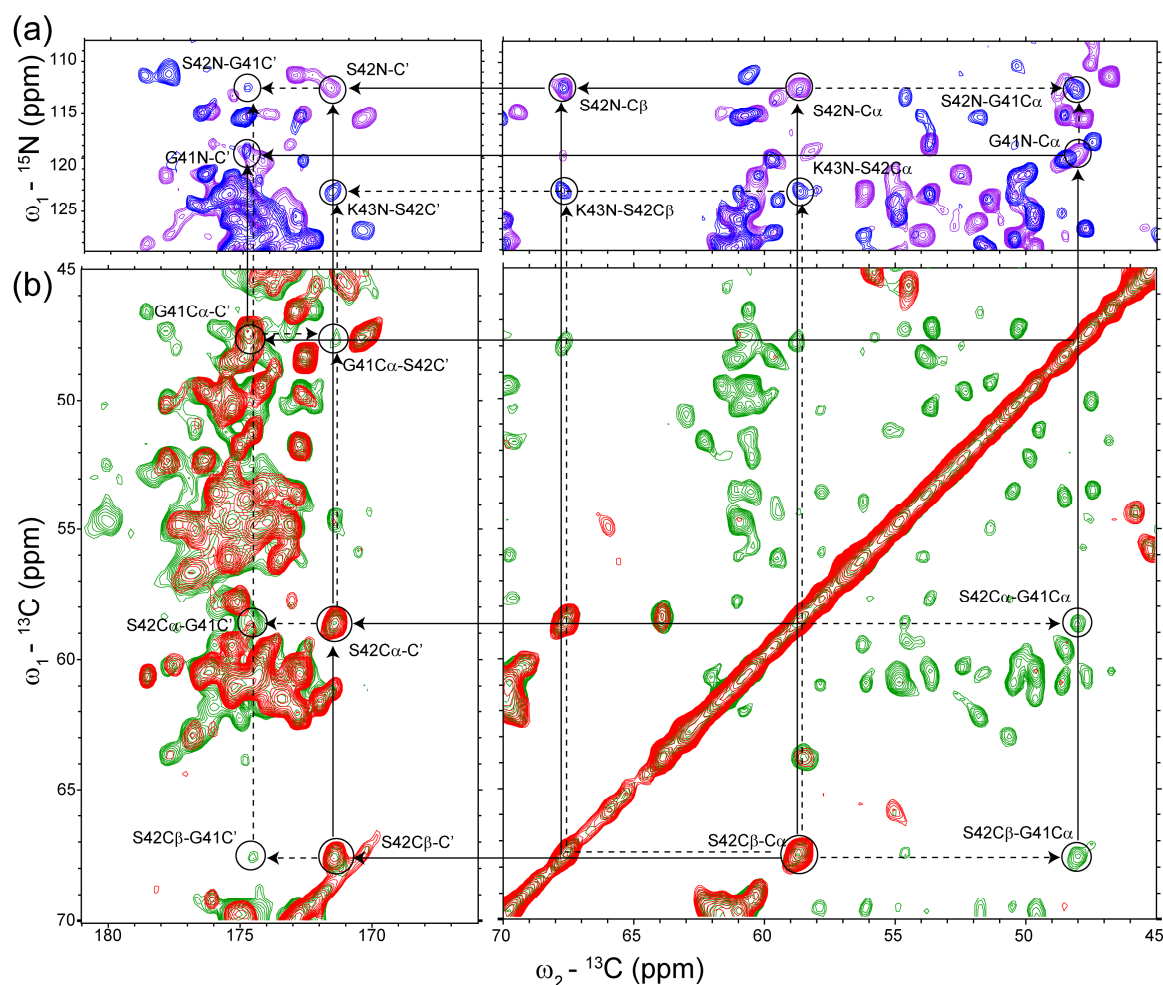


Figure 3.3. Sequential resonance assignment strategy using 2D ^{15}N - ^{13}C and ^{13}C - ^{13}C correlation spectra recorded on $[\text{U-}^{13}\text{C}]$ Glc-labeled $m\alpha\text{S}$ fibrils. Excerpts of carbonyl (left) and aliphatic (right) regions of (a) NCACX (purple) and NCOCX (blue) spectra, and (b) ^{13}C - ^{13}C PSDS spectra with mixing times of 20 ms (red) and 150 ms (green). Intra- and inter-residue connectivities are illustrated with solid and dashed lines, respectively. The spectra in (a) were recorded at 18.8 Tesla (^1H resonance frequency: 800 MHz), and the spectra in (b) were recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz). All the spectra were measured at a spinning frequency of 11 kHz. All the spectra were zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

To better resolve and assign carbons with similar chemical shifts, which appeared close to the diagonal or were overlapped with the diagonal of the 2D ^{13}C - ^{13}C PSDS spectra, we recorded a 2D (2Q,1Q) correlation spectrum with the SPC5 scheme [98]. E.g., the chemical shifts of $^{13}\text{C}\beta$ and $^{13}\text{C}\gamma$ of Gln79 are so close that the correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$ and $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ were overlapped into one strong cross peak, and the correlation of $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma$ was overlapped with the diagonal in the 2D ^{13}C - ^{13}C PSDS spectrum (black

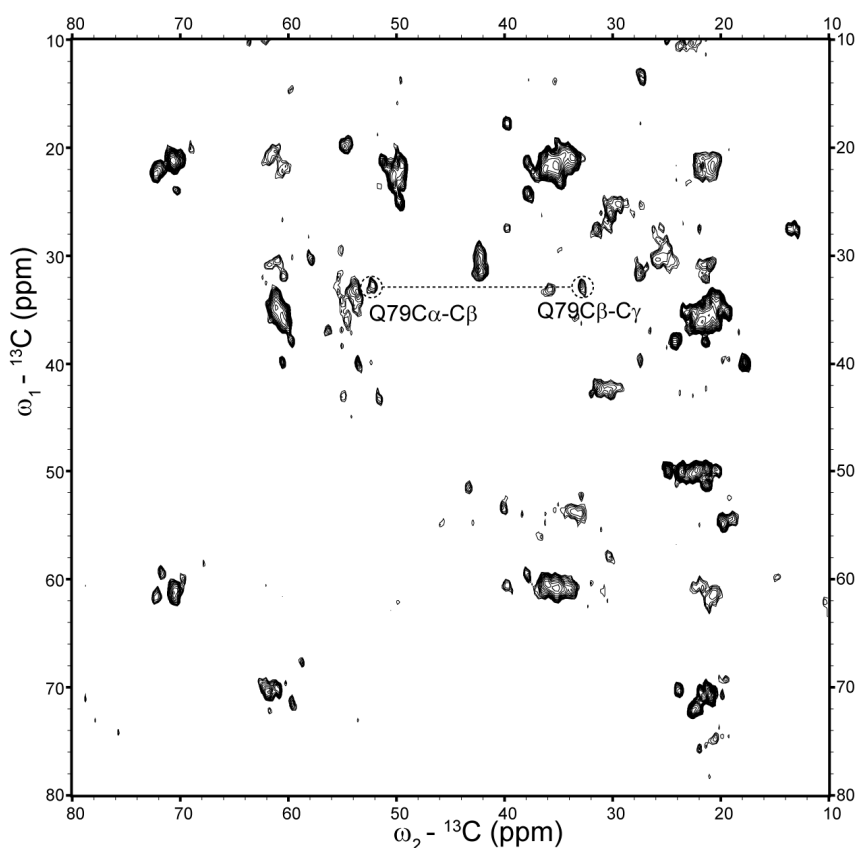


Figure 3.4. 2D (2Q,1Q) correlation spectrum with the SPC5 scheme recorded on [U- ^{13}C]Glc-labeled αS fibrils. Both spectra were recorded at 18.8 Tesla (^1H resonance frequency: 800 MHz) and at a spinning frequency of 8 kHz. The spectrum was zero-filled to 1024 points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

spectrum in Fig. 3.2). However, with the recorded 2D (2Q,1Q) correlation spectrum, as shown in Fig. 3.4, the one-bond correlation of Gln79 $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma$ was well revealed without the disturbance of the diagonal region and was unambiguously assigned. Additionally, the 2D (2Q,1Q) correlation spectrum only contains one-bond correlations due to its dependence of 2Q coherences on strong dipolar couplings. Thereby, only the one-bond correlation of Gln79 $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$ is predicted to be present and the unambiguous assignment was thus obtained, while the two-bond Gln79 $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ should not be observed.

Additionally, 3D NCACX and NCOCX spectra were recorded to confirm and obtain more unambiguous assignments (Fig. 3.5). The spectral overlap observed in the 2D ^{13}C - ^{15}N and ^{13}C - ^{13}C correlation spectra was substantially alleviated by adding an additional dimension. Basically, the 3D NCACX experiment establishes intra-residue

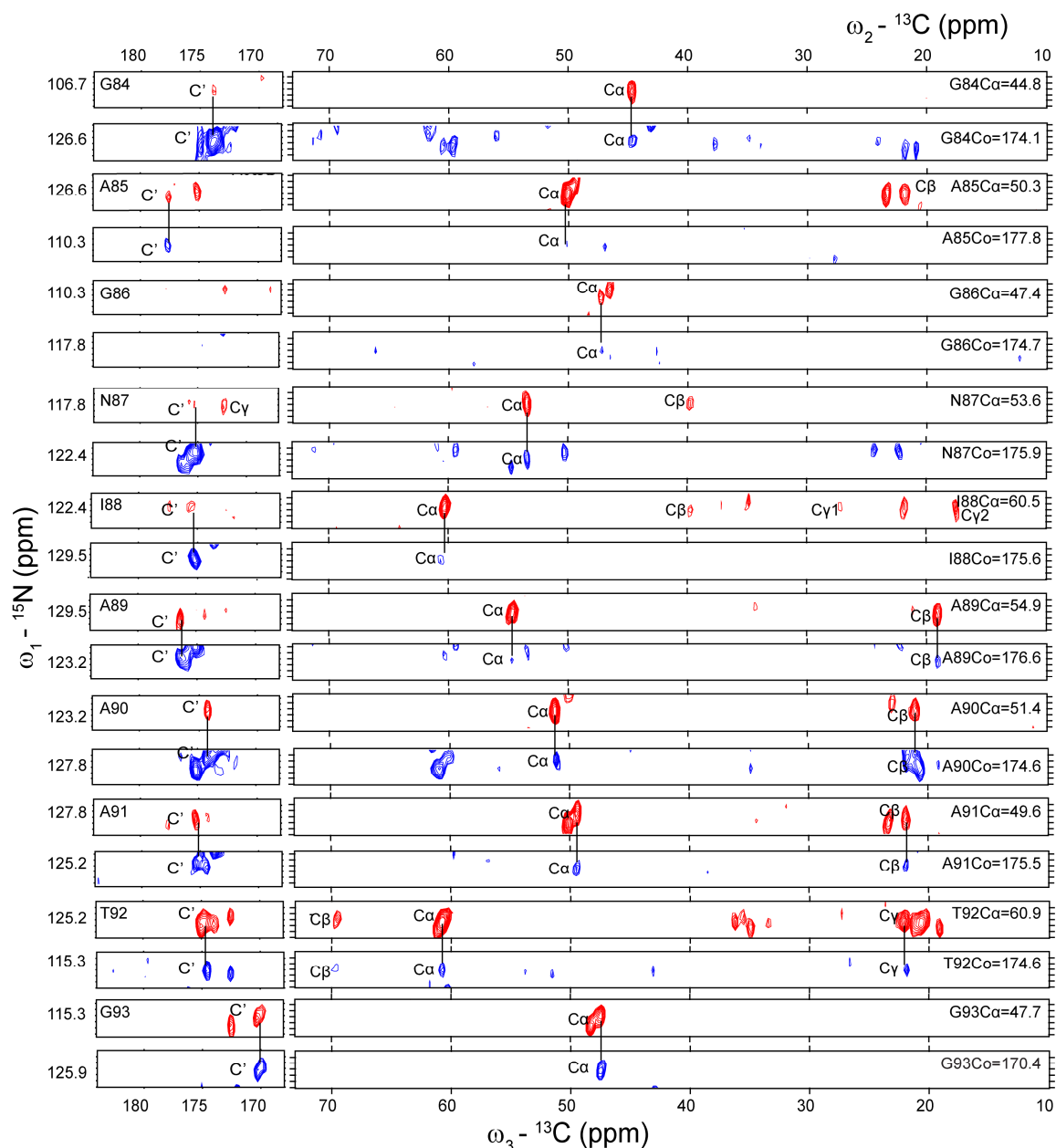


Figure 3.5. Strip plots from 3D NCACX (red) and 3D NCOCX (blue) spectra of $[\text{U-}^{13}\text{C}]$ Glc-labeled $\text{m}\alpha\text{S}$ fibrils, showing the sequential resonance assignment of Gly84-Gly93. Both spectra were recorded at 20.0 Tesla (${}^1\text{H}$ resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. The 3D NCACX spectrum was zero-filled to 1024 (F3, direct dimension), and 64 (F1 and F2, indirect dimensions) points, and apodized with a squared sine bell window function (51.4° shift) along three dimensions. The 3D NCOCX spectrum was zero-filled to 1024 (F3, direct dimension), and 64 (F1 and F2, indirect dimensions) points, and apodized with a squared sine bell window function (60° shift) along three dimensions.

backbone-sidechain correlations (i.e., ${}^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)\text{-}^{13}\text{C}\text{X}(i)$), whereas 3D NCOCX relates the backbone shift of ${}^{15}\text{N}(i)$ to the carbon shifts of the preceding residue (i.e., ${}^{15}\text{N}(i)\text{-}^{13}\text{C}'(i-1)\text{-}^{13}\text{C}\text{X}(i-1)$). As an example, the correlation of ${}^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)\text{-}^{13}\text{C}\alpha(i)$ at 106.7-44.8-44.8 ppm in the 3D NCACX spectrum (red spectrum in Fig. 3.5) was

identified to be the spin system of Gly, and the chemical shift of $^{15}\text{N}(i+1)$ was then identified to be 126.6 ppm in the 3D NCOCX spectrum (blue spectrum in Fig. 3.5). Subsequently, the $(i+1)$ residue was identified to be Ala (i.e., 50.3-21.9 ppm) in the 3D NCACX, according to its characteristic carbon chemical shifts. As an example, the sequential resonance assignment of Gly84-Gly93 is shown in Fig. 3.5. Similarly, more sequential resonance assignments could be obtained using the same strategy.

Taken together, the combination of 2D and 3D ^{13}C - ^{15}N and ^{13}C - ^{13}C spectra of [U- $^{13}\text{C}/^{15}\text{N}$]-labeled αS fibrils (Table A1 in appendix A) enabled us to obtain the sequential connections for the following stretches unambiguously: Gly41-Val48, Ala56-Lys60, Asn65-Val70, and Thr72-Phe94. In total, 78% of residues from Gly41 to Phe94 were assigned, including the backbone and most sidechain resonances.

3.3 Sparse labeling schemes: [2-¹³C]glucose and [1-¹³C]glucose

The limited resolution and spectral overlap observed for the [U-¹³C]Glc-labeled samples are major obstacles for obtaining the complete sequential resonance assignment. Particularly, the full sequential resonance assignment for m α S fibrils is complicated, not only due to the highly repetitive, imperfect KTKEGV motifs (Fig. 1.4), but also due to the similar secondary structure adopted by most residues in the β -sheet arrangement. Another difficulty to complete the full assignment of m α S fibrils relates to the high abundance and sequential clustering of Val and Thr, e.g., Val48-Val49, Val52-Thr53-Thr54-Val55, Val70-Val71-Thr72, Val74-Thr75, Thr81-Val82, since the C α and C γ chemical shifts of Val and Thr are similar. In order to alleviate spectral overlap and enhance the resolution, we used sparse ¹³C labeling schemes based on protein expression in minimal medium containing [2-¹³C]Glc or [1-¹³C]Glc as the sole carbon source. [2-¹³C]Glc and [1-¹³C]Glc are carbon sources that are isotopically ¹³C labeled only on one out of six carbons, while [2-¹³C]glycerol and [1,3-¹³C]glycerol are carbon sources that are isotopically ¹³C labeled on one out of three carbons and two out of three carbons, respectively. Therefore, bacterial growth medium containing either [2-¹³C]Glc or [1-¹³C]Glc as the sole carbon source produces amino acids in an ultra-sparse ¹³C enrichment pattern, where only one out of six carbons is ¹³C labeled. With the ultra-sparse labeling schemes, not only the spectral resolution is enhanced substantially due to the removal of the majority of one-bond ¹³C-¹³C J and dipolar couplings, but also the resulting spectrum is simplified dramatically by a reduced number of cross peaks. Moreover, the sequential transfer, e.g., ¹³C α (i)-¹³C α (i \pm 1), is facilitated by the ¹³C spin dilution. Potentially, these two sparsely labeling schemes will facilitate collection of long-range distance restraints.

3.3.1 Labeling pattern of [2-¹³C]Glc labeling scheme

Briefly, three major metabolic pathways are involved in biosynthesis of the 20 standard amino acids in *E. coli*, glycolysis, PPP, and TCA cycle [100]. The three major metabolic pathways involved in the biosynthesis of amino acids using [2-¹³C]Glc as the sole carbon source are outlined in Fig. 3.6. Individual biosynthetic pathways of the 20 amino acids using [2-¹³C]Glc are shown in Fig. A7-8. According to their biosynthetic precursors, we categorize the 20 amino acids into 3 groups, a) group I, amino acids derived from glycolytic intermediates, i.e., Ser, Gly, Cys, Ala, Val and Leu, b) group II,

amino acids produced from PPP intermediates and glycolytic intermediates, i.e., Trp, Phe, Tyr, and His, and c) group III, amino acids synthesized from TCA cycle intermediates, i.e., Asn, Asp, Ile, Lys, Met, Thr, Glu, Gln, Arg and Pro.

The precursors for group I amino acids of a) Ser, Gly, and Cys, and b) Ala, Val and Leu are a) 3PG, and b) pyruvate, respectively. 3PG and pyruvate are generated during glycolysis, and approximately 1/2 ^{13}C enrichment at position 2 is resulted from [2- ^{13}C]Glc, since the carbons 2 and 5 of glucose contribute equally and indistinguishable to carbon 2 of 3PG and pyruvate (Fig. 3.6a). As a result, about 1/2 of the $\text{C}\alpha$ is ^{13}C enriched for the group I amino acids without simultaneously being ^{13}C labeled at the $\text{C}\beta$ position except for Val and Leu. For Val, the isotopomers of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$ (1/4), $^{13}\text{C}\alpha$ - $^{12}\text{C}\beta$ (1/4), $^{12}\text{C}\alpha$ - $^{13}\text{C}\beta$ (1/4), and $^{12}\text{C}\alpha$ - $^{12}\text{C}\beta$ (1/4) are formed from 1/2 [2- ^{13}C]pyruvate and 1/2 [^{12}C]pyruvate (Fig. 3.7a). For Leu, the isotopomers of $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma$ - $^{13}\text{C}'$ (1/8), $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma$ - $^{12}\text{C}'$ (1/8), $^{13}\text{C}\beta$ - $^{12}\text{C}\gamma$ - $^{13}\text{C}'$ (1/8), $^{12}\text{C}\beta$ - $^{13}\text{C}\gamma$ - $^{13}\text{C}'$ (1/8), $^{13}\text{C}\beta$ - $^{12}\text{C}\gamma$ - $^{12}\text{C}'$ (1/8), $^{12}\text{C}\beta$ - $^{13}\text{C}\gamma$ - $^{12}\text{C}'$ (1/8), and $^{12}\text{C}\beta$ - $^{12}\text{C}\gamma$ - $^{12}\text{C}'$ (1/8) are generated from 1/2 [2- ^{13}C]pyruvate and 1/2 [^{12}C]pyruvate, and $\text{C}\alpha$ is predicted not to be ^{13}C labeled (Fig. 3.7a). As an illustration, the labeling patterns together with their detailed isotopomers of group I amino acids are shown in Fig. 3.7.

The precursors for group II amino acids of a) Trp, b) Tyr and Phe, and c) His are a) 3PG, PEP, E4P and R5P, b) PEP and E4P, and c) PEP and R5P, respectively. Due to the fact that these aromatic amino acids are biosynthesized via more elaborate pathways, their detailed biosynthetic pathways are omitted here. To guide the eyes for their ultimate carbon sources, individual carbons of their backbone and sidechain carbons are indicated with arrows in Fig. A7c (appendix A). $^{13}\text{C}\alpha$ / $^{13}\text{C}\gamma$ of Phe and Tyr, and $^{13}\text{C}\alpha$ / $^{13}\text{C}\delta 2$ of Trp are derived from the carbon 2 of PEP and 3PG, an enrichment level of about 1/2 for these ^{13}C spins is thus expected. The other ^{13}C carbons for the group III amino acids are derived from E4P and/or R5P. Both E4P and R5P are biosynthesized in the PPP. As shown in Fig. 3.6b, E4P will be ^{13}C labeled at positions 1 and/or 3, and R5P will be ^{13}C labeled at positions 2 and/or 4 or separately at position 1, due to the rapid equilibration of the transketolase and pentose-5-phosphate isomerization reactions [101; 102; 103]. As a result, spin pair $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ or a single spin $^{13}\text{C}\delta 2$ for His are formed from [2,4- ^{13}C]R5P or

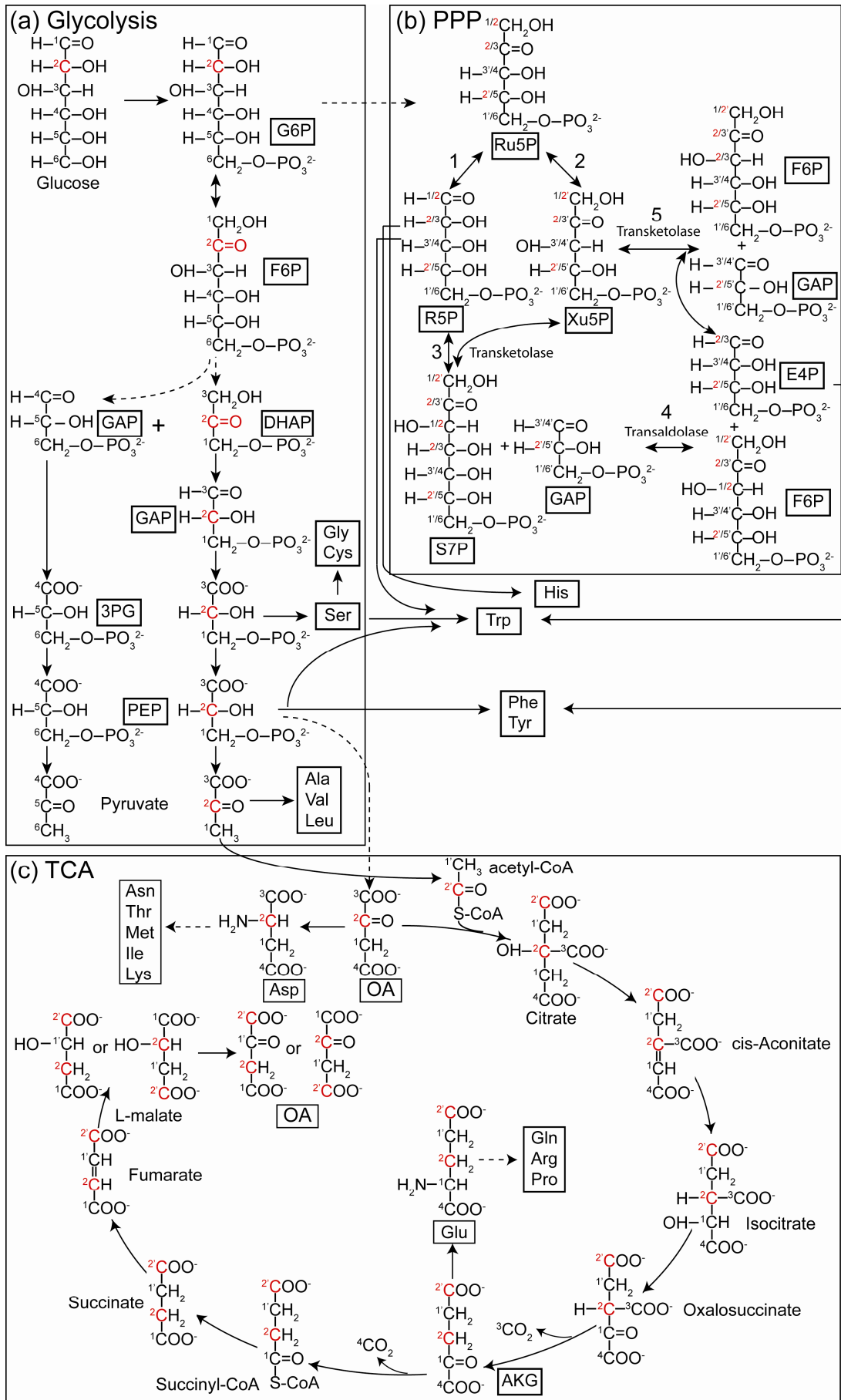


Figure 3.6. Key metabolic pathways involved in the biosynthesis of the 20 standard amino acids when *E. coli* is grown in a minimal medium containing [2-¹³C]Glc as the sole carbon source, (a) glycolysis, (b) PPP, and (c) TCA cycle. The *superscript numbers* do not correspond to the nomenclature number of a given molecule but indicate the original carbon position in the six-carbon [2-¹³C]Glc chain. The carbons that are ¹³C labeled are colored in red. For the sake of clarity, sugars are shown in their linear forms. For simplicity, some steps are not included, which are indicated by dashed arrows. Double-headed arrows indicate reversible reactions. Abbreviations: Pentose phosphate pathway, PPP; tricarboxylic acid cycle, TCA cycle; Glucose-6-phosphate, G6P; Fructose-6-phosphate, F6P; Dihydroxyacetone phosphate, DHAP; Glyceraldehyde-3-phosphate, GAP; 3-phosphoglycerate, 3PG; phosphoenolpyruvate, PEP; Ribulose-5-phosphate, Ru5P; Ribose-5-phosphate, R5P; Xylulose-5-phosphate, Xu5P; Erythrose-4-phosphate, E4P. Sedoheptulose-7-phosphate, S7P; Oxaloacetate, OA; α -Ketoglutarate, AKG.

[1-¹³C]R5P, respectively. For Phe and Tyr, ¹³C ϵ 1-¹³C ϵ 2 are formed from [1,3-¹³C]E4P. For Trp, ¹³C γ is formed from [2,4-¹³C]R5P or ¹³C δ 1 is formed from [1-¹³C]R5P, and ¹³C ζ 2-¹³C ζ 3 are formed from [1,3-¹³C]E4P. However, the enrichment level of these ¹³C labeled atoms could not be determined, instead, a 1/2 distribution is displayed in Fig. 3.7b.

The precursors of group III amino acids of a) Asn, Asp, Ile, Lys, Met, and Thr, and b) Glu, Gln, Arg and Pro, are a) OA, and b) AKG, respectively. Following the steps in Fig. 3.6c, pyruvate is oxidized to acetyl-CoA, and enters the TCA cycle, where the two key intermediates OA and AKG are formed. OA also derives directly from PEP (Fig. 3.6) through several anaplerotic reactions. In principle, the labeling pattern for the group III amino acids is complicated due to the cyclic property of TCA cycle, which spreads the single label in the starting compound to multiple carbon sites. Multiple isotopomers for the group III amino acids are thus formed, and the population of individual isotopomers for the 10 amino acids varies upon the passes of TCA cycle occur. In total, 6 types of OA (with at least one ¹³C label) are yielded in TCA cycle by using [2-¹³C]Glc as the sole carbon source in the protein expression medium, [2-¹³C]OA, [1,3-¹³C]OA, [2,4-¹³C]OA, [1,4-¹³C]OA, [1-¹³C]OA, and [4-¹³C]OA (Fig. A8b). According to the biosynthetic pathways (Fig. A8a), 6 isotopomers (with at least one ¹³C label) are thus expected for OA derived amino acids with the exceptions of Ile and Lys. The same labeling pattern is expected for Asp, Asn, Thr, and Met, and the following ¹³C labeled single spins or spin pairs are expected, ¹³C α , ¹³C α -¹³C γ , ¹³C β -¹³C γ' , ¹³C γ -¹³C γ' , ¹³C γ , and ¹³C γ' (Fig. 3.8). Consequently, the enrichment level of each isotopomer is estimated to be 1/12 for Asp, Asn, Thr, and Met (i.e., 1/6 \times 1/2). The labeling patterns of Ile and Lys are more complicated. As shown in Fig. A8a, Ile is converted from a molecule of Thr and a molecule of pyruvate, and Lys is converted from a molecule of Asp and a molecule of

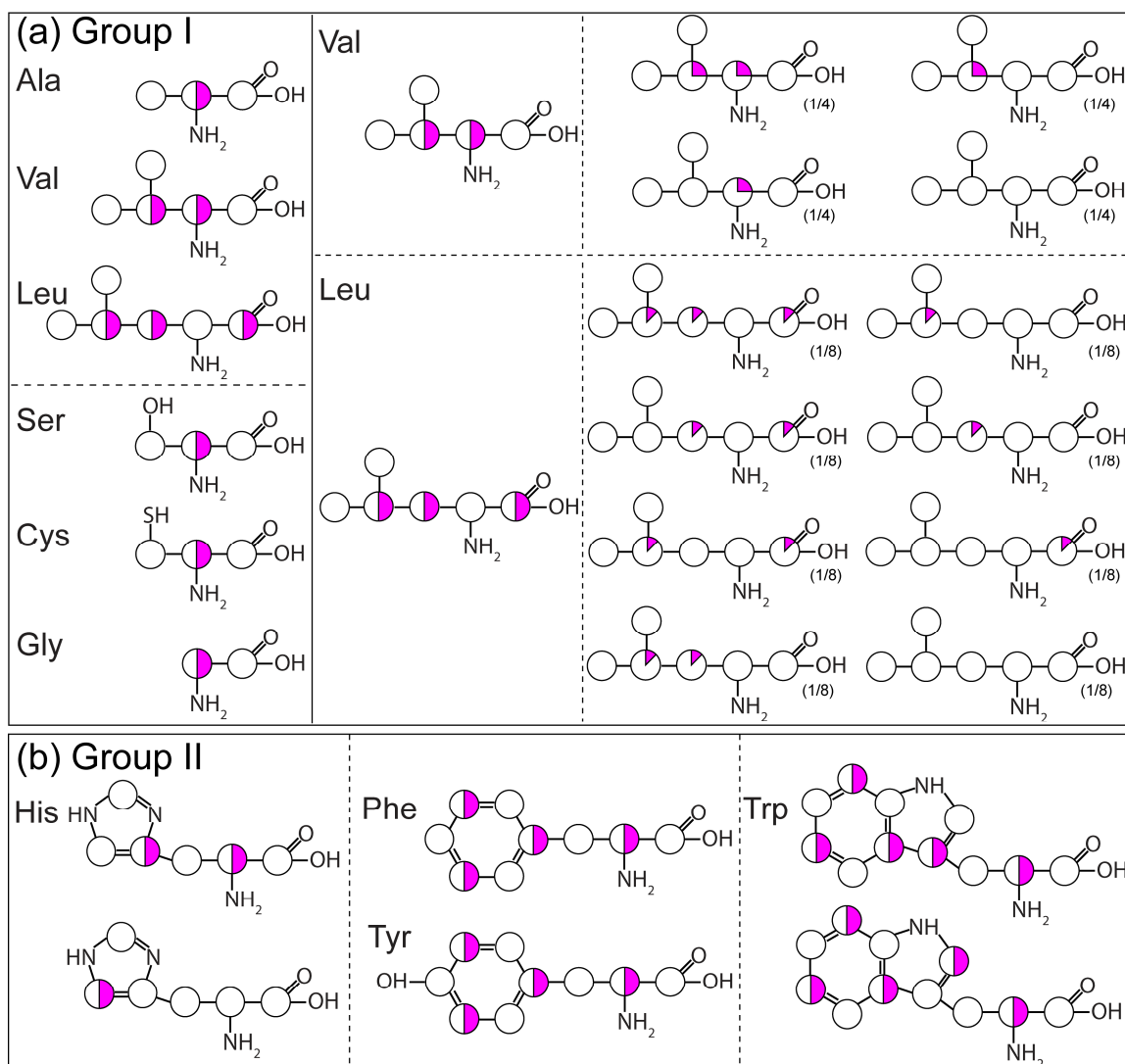


Figure 3.7. Schematic representation of the labeling patterns for amino acids synthesized from (a) the glycolytic intermediates (i.e., group I amino acids) and (b) the PPP intermediates as well as the glycolytic intermediates (i.e., group II amino acids) in a $[2-^{13}\text{C}]\text{Glc}$ containing medium. Carbons that are ^{13}C labeled are filled with magenta color, and the enrichment level is indicated by the percentage of the filled color. The percentage of the ^{13}C spins derived from E4P or R5P, i.e., His $^{13}\text{C}\alpha/^{13}\text{C}\gamma/^{13}\text{C}\delta 2$, Phe/Tyr $^{13}\text{C}\epsilon 1/^{13}\text{C}\epsilon 2$, and Trp $^{13}\text{C}\gamma/^{13}\text{C}\delta 1/^{13}\text{C}\zeta 1/^{13}\text{C}\zeta 2$, could not be determined, a 1/2 distribution is instead displayed, and their isotopomers are not considered here. Val and Leu are the exceptions of group I amino acids, and the detailed isotopomers are displayed in right. For clarity, the percentage of each isotopomer of Val and Leu is shown below.

pyruvate. For Ile, $\text{C}\beta$ derives from carbon 2 of pyruvate, the possibility of being ^{13}C labeled is thus about 1/2. The ^{13}C labeled single spins or spin pairs of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$, $^{13}\text{C}\alpha$, $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta 1$ - $^{13}\text{C}\beta$, $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta 1$, $^{13}\text{C}\gamma 1$ - $^{13}\text{C}'$ - $^{13}\text{C}\beta$, $^{13}\text{C}\gamma 1$ - $^{13}\text{C}'$, $^{13}\text{C}\delta 1$ - $^{13}\text{C}'$ - $^{13}\text{C}\beta$, $^{13}\text{C}\delta 1$ - $^{13}\text{C}'$, $^{13}\text{C}\delta 1$ - $^{13}\text{C}\beta$, $^{13}\text{C}\delta 1$, $^{13}\text{C}'$ - $^{13}\text{C}\beta$, $^{13}\text{C}'$, and $^{13}\text{C}\beta$ are formed for Ile. The relative enrichment level of individual isotopomers for Ile is indicated in Fig. 3.8. From the biosynthetic pathway of Lys, tracing of isotopes through the final conversion is even more

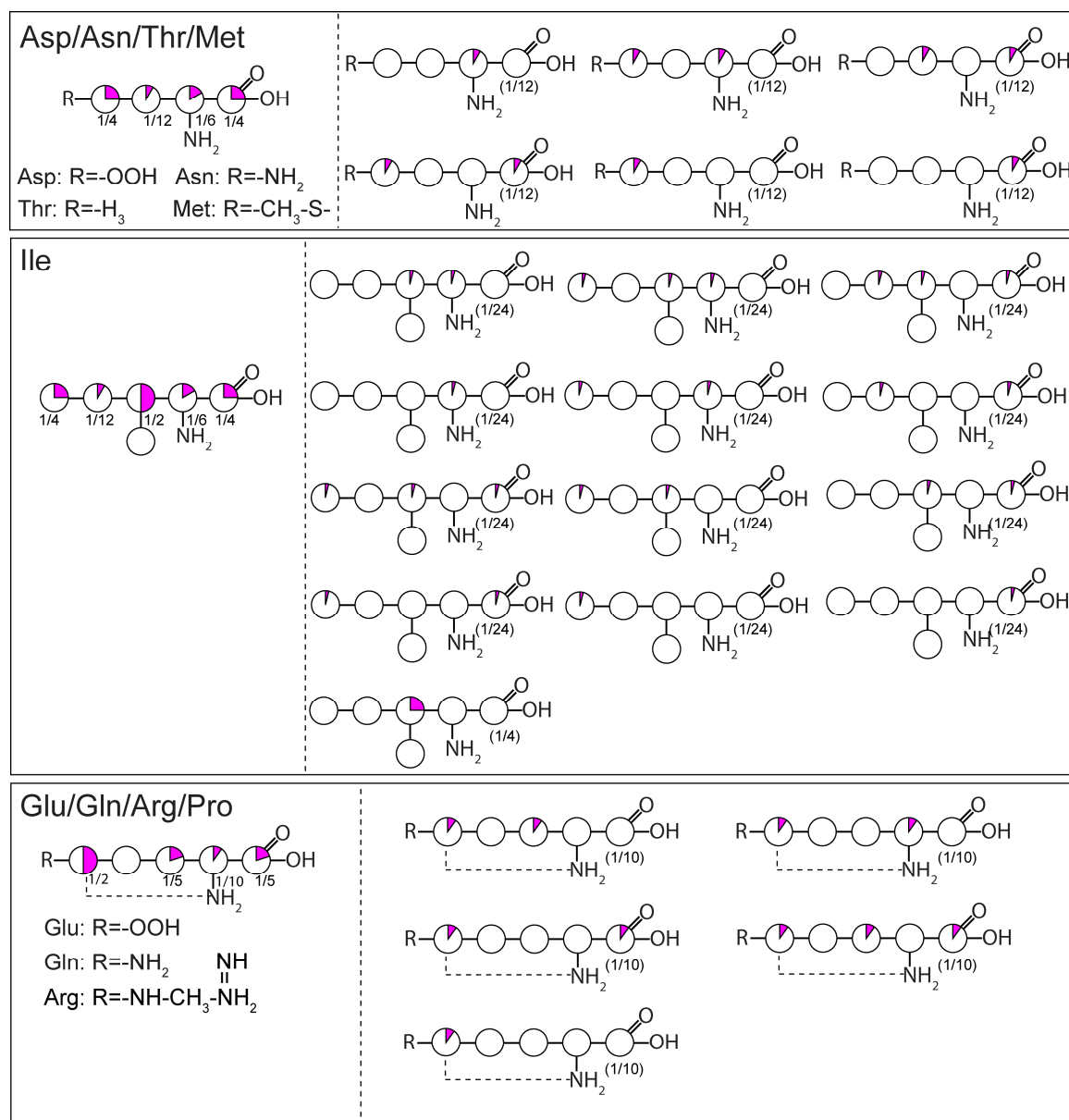


Figure 3.8. Schematic representation of the expected labeling pattern for group III amino acids produced in a [2-¹³C]Glc containing medium for protein expression. Carbons that are ¹³C labeled are filled with magenta color, and the enrichment level is indicated by the percentage of the filled color. For clarity, only the isotopomers with at least one ¹³C label are shown. The list of isotopomers for Lys is shown in Fig. 3.9.

complicated due to the occurrence of the symmetric intermediate DAP (Fig. A8a). Carbons 2 and 6 of DAP are equivalent, leading to a scrambling of isotope label on the pathway to Lys. The ¹³C labeled single spins, spin pairs or spin triples of a) ¹³C α -¹³C ϵ , ¹³C α , ¹³C α -¹³C γ -¹³C ϵ , ¹³C α -¹³C γ , ¹³C β -¹³C', ¹³C β -¹³C', ¹³C γ -¹³C', ¹³C γ -¹³C', ¹³C γ -¹³C ϵ , ¹³C γ , ¹³C', ¹³C', and ¹³C ϵ are formed if C α derives from carbon 2 of OA, and b) ¹³C ϵ -¹³C α , ¹³C ϵ , ¹³C ϵ -¹³C γ -¹³C α , ¹³C ϵ -¹³C γ , ¹³C δ -¹³C α , ¹³C δ , ¹³C γ -¹³C α , ¹³C γ , and

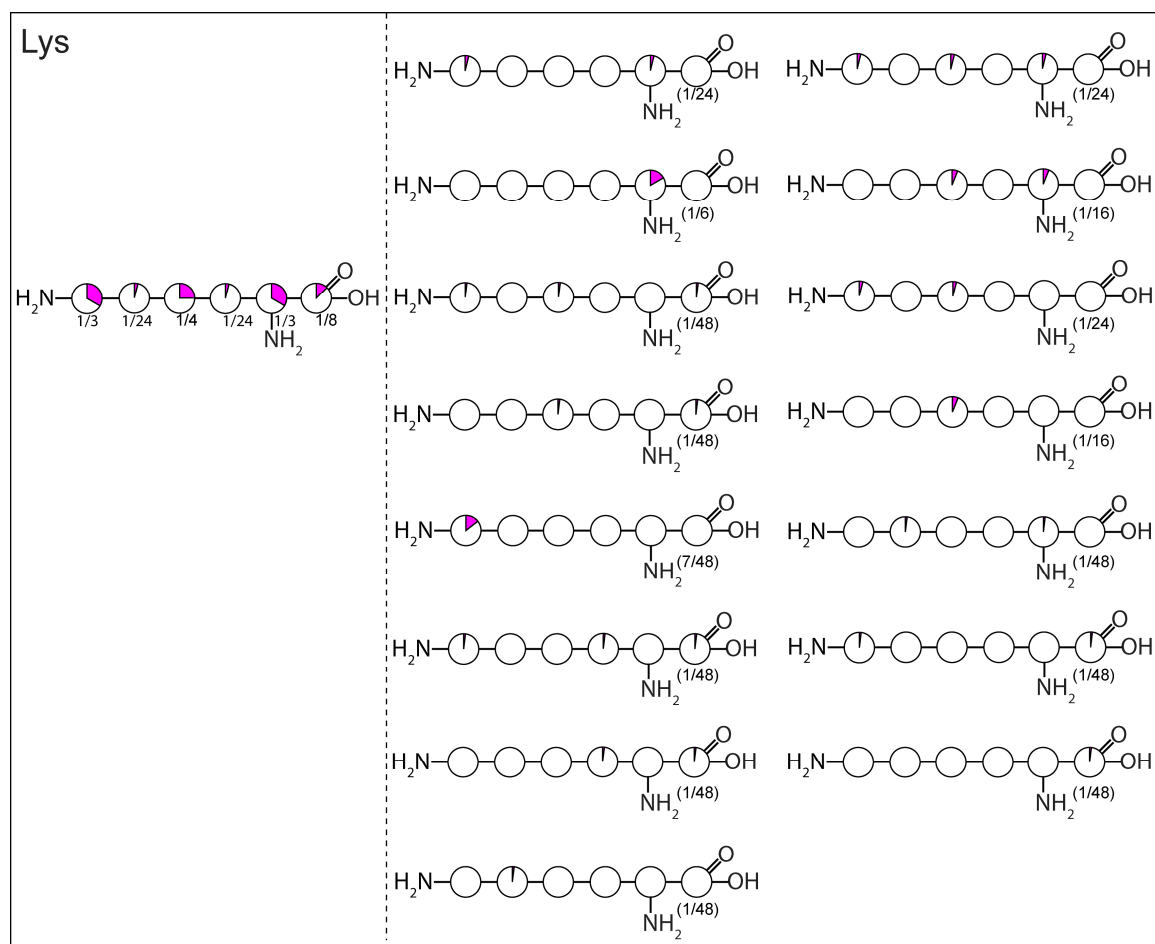


Figure 3.9. The isotopomers of Lys produced in a $[2-^{13}\text{C}]$ Glc containing medium for protein expression. Carbons that are ^{13}C labeled are filled with magenta color, and the enrichment level is indicated by the percentage of the filled color. For clarity, only the isotopomers with at least one ^{13}C label are shown.

$^{13}\text{C}\alpha$ are formed if $\text{C}\alpha$ derives from carbon 2 of pyruvate. In total, 15 isotopomers with ^{13}C labels are expected for Lys (Fig. 3.9). The relative enrichment level of the expected isotopomers is indicated in Fig. 3.8 based on the assumption that $1/2 \text{ C}\alpha$ is derived from the carbon 2 of OA and $1/2 \text{ C}\alpha$ is derived from the carbon 2 of pyruvate.

Likewise, the labeling pattern of the AKG derived amino acids also depends on the number of the rounds of the TCA cycle. As shown in Fig. A8d, 5 isotopomers of AKG (with at least one ^{13}C label) are obtained in the TCA cycle, $[3,5-^{13}\text{C}]\text{AKG}$, $[2,5-^{13}\text{C}]\text{AKG}$, $[1,3,5-^{13}\text{C}]\text{AKG}$, $[1,5-^{13}\text{C}]\text{AKG}$, and $[5-^{13}\text{C}]\text{AKG}$. As a consequence, 5 isotopomers with alternating ^{13}C labels (i.e., $^{12}\text{C}-^{13}\text{C}-^{12}\text{C}$ or $^{13}\text{C}-^{12}\text{C}-^{13}\text{C}$) are thus formed for Glu, Gln, Arg and Pro, i.e., $^{13}\text{C}\beta-^{13}\text{C}\delta$, $^{13}\text{C}\alpha-^{13}\text{C}\delta$, $^{13}\text{C}'-^{13}\text{C}\beta-^{13}\text{C}\delta$, $^{13}\text{C}'-^{13}\text{C}\delta$, and $^{13}\text{C}\delta$. The enrichment level of each isotopomer for the AKG derived amino acids is estimated to be about $1/10$ (i.e., $1/5 \times 1/2$, Fig. 3.8).

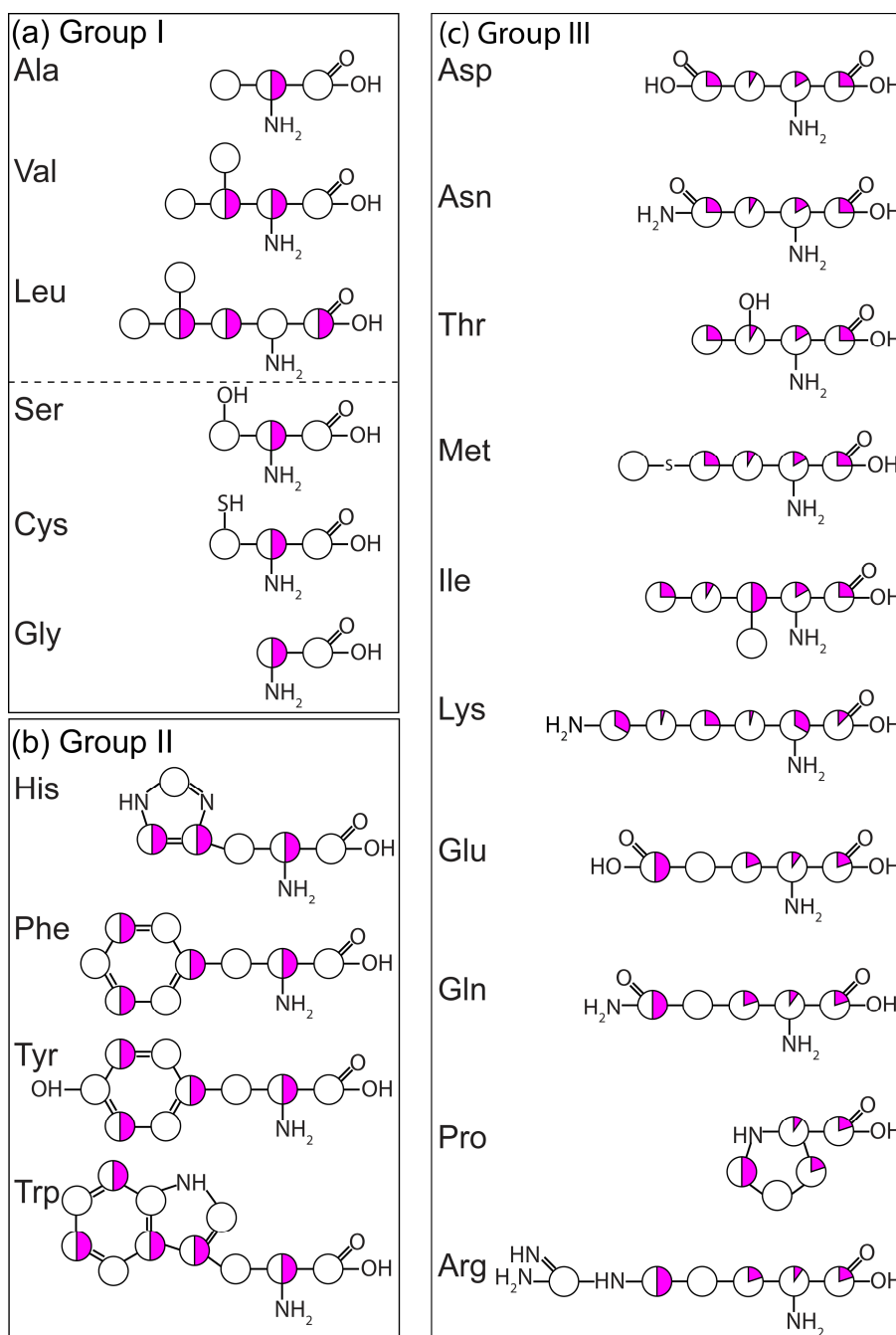


Figure 3.10. Schematic representation of the expected labeling pattern for the 20 amino acids produced in a $[2-^{13}\text{C}]\text{Glc}$ containing medium for protein expression. (a) Group I, (b) group II, and (c) group III amino acids. Carbons that are ^{13}C labeled are filled with magenta color, and the enrichment level is indicated by the percentage of the filled color. For the amino acids that have multiple labeling patterns, the detailed isotopomers are displayed in Fig. 3.7-3.9.

However, we need to note that the ^{13}C enrichment levels at $\text{C}\alpha$ position for OA (with the exception of Lys) and AKG derived amino acids using the $[2-^{13}\text{C}]\text{Glc}$ labeling scheme measured by solution NMR are $28 \pm 2\%$ and $17 \pm 2\%$ [61], respectively. As mentioned above, the labeling patterns and enrichment levels for the group III amino

acids vary upon passes of the TCA cycle occur. In general, small numbers of passes through the TCA cycle minimize scrambling of ^{13}C labels, thus increasing the percentage of the isotopomers derived from small numbers of passes of TCA cycle. The enrichment level of each isotopomer for the group III amino acids is estimated based on the types of isotopomers, i.e., more than 3 and 4 passes of the TCA cycle are assumed to occur for OA and AKG derived amino acids, respectively. It is thus reasonable that the estimated enrichment level of $^{13}\text{C}\alpha$ for the group III amino acids is relatively lower than the isotopomers (with at least one ^{13}C spin) with a relative labeling percentage. Importantly, the illustration of individual isotopomers (with at least one ^{13}C spin) reveals that, all the group III amino acids except Ile are in an alternating ^{13}C labeling pattern (Fig. 3.8 and 3.9), i.e., no one-bond ^{13}C - ^{13}C couplings.

Taken together, according to the above analysis, we propose a new labeling pattern sheet for all the 20 standard amino acids produced in an *E.coli* medium containing $[2\text{-}^{13}\text{C}]\text{Glc}$ as the sole carbon source for protein expression (Fig. 3.10). In total, $^{13}\text{C}\alpha$ is labeled for 17 out of 20 amino acids without being simultaneously ^{13}C labeling at the $\text{C}\beta$ (with the exceptions of Val, Leu and Ile). Importantly, no ^{13}C is enriched at C' simultaneously with the $^{13}\text{C}\alpha$ (Fig. 3.7-3.9). In other words, an alternating labeling pattern is achieved for 17 out of 20 amino acids using $[2\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme, i.e., either ^{12}C - ^{13}C - ^{12}C or ^{13}C - ^{12}C - ^{13}C .

3.3.2 Resolution enhancement and spectrum simplification using $[2\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme

With this ultra-sparse labeling scheme based on $[2\text{-}^{13}\text{C}]\text{Glc}$, the resolution of the ssNMR spectra is enhanced substantially by removing the majority of one-bond ^{13}C - ^{13}C J and dipolar couplings. In addition, the spectral overlap is alleviated dramatically by a reduced number of cross peaks. The knowledge of the labeling pattern, especially the expected ^{13}C labeled spin pairs (Fig. 3.7-3.9), affords valuable information for a rapid identification of spin systems. As an illustration, we recorded a 2D ^{13}C - ^{13}C PDSO spectrum with a short mixing time of 100 ms on $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled m α S fibrils. With the short mixing time, the polarization is mainly stored on the intra-residue ^{13}C spins. Consequently, a rapid identification of Val, Leu, Asp, Asn, Met, Thr, Ile, Lys, Glu, Gln, Arg, Pro, Phe, Tyr, Trp, and His by intra-residue ^{13}C - ^{13}C correlations is expected to be

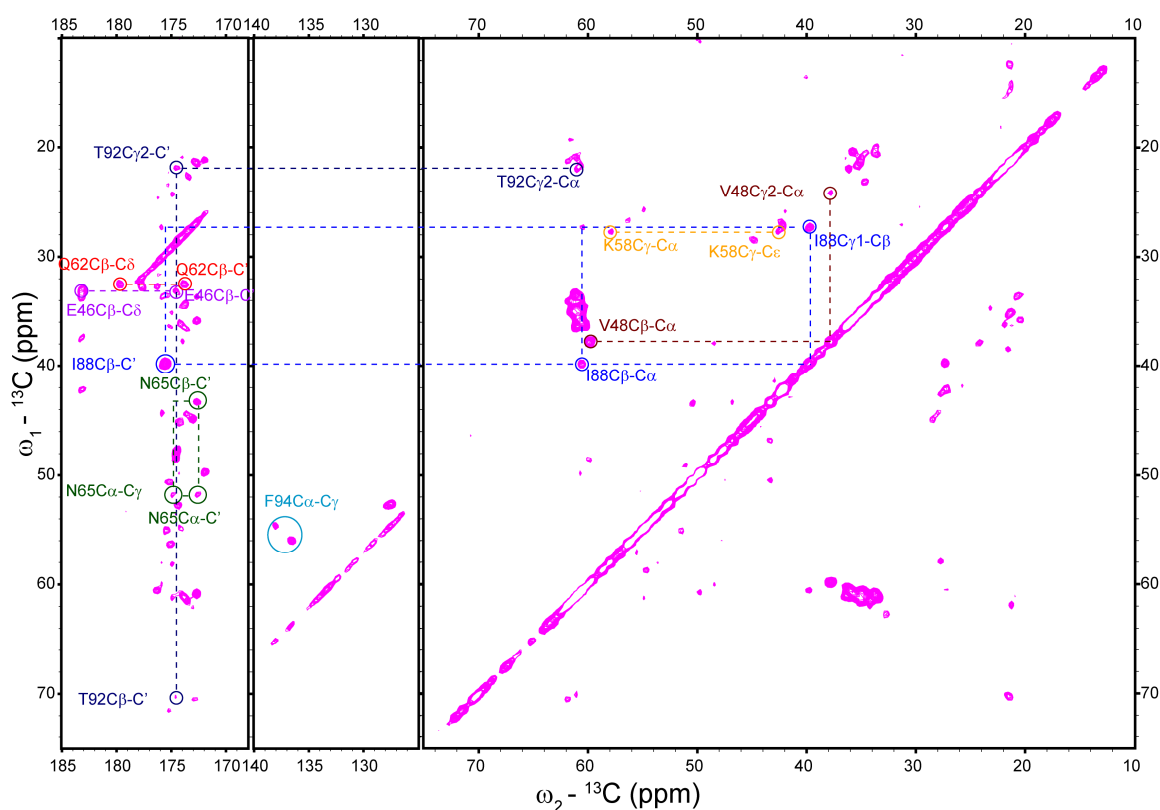


Figure 3.11. 2D ^{13}C - ^{13}C PDSM spectrum with a mixing time of 100 ms on $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils. Correlations of residue Glu46, Val48, Lys58, Gln62, Asn65, Ile88, and Thr92 for $\text{m}\alpha\text{S}$ are illustrated with dashed lines, whereas cross peaks are indicated with circles, otherwise, no cross peak belongs to this spin system. Correlations of Phe94 $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ are circled in blue. The spectrum was recorded at 14.1 Tesla (^1H resonance frequency: 600 MHz) and at a spinning frequency of 11 kHz. The spectrum was zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

possible given the information about the characteristic chemical shifts of the amino acids. Trp and Arg are not present in the primary sequence of $\text{m}\alpha\text{S}$ (Fig. 1.4). The unambiguous assignment of Leu, Asp, Pro, Tyr, Met, and His is not obtained so far, probably because these residues are either in or close to the flexible part of $\text{m}\alpha\text{S}$ fibrils (see the discussion in section 3.4 of Part I).

Taking Thr as an example for OA derived amino acids except Lys and Ile (i.e., Asp, Asn, Met, and Thr), two-bond correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma_2$ and $^{13}\text{C}\beta$ - $^{13}\text{C}'$, and three-bond correlations of $^{13}\text{C}\gamma_2$ - $^{13}\text{C}'$ are expected (Fig. 3.8). Indeed, the correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma_2$, $^{13}\text{C}\beta$ - $^{13}\text{C}'$, and $^{13}\text{C}\gamma_2$ - $^{13}\text{C}'$ were observed for Thr92, while the unexpected correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$, $^{13}\text{C}\alpha$ - $^{13}\text{C}'$ and $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma_2$ were not observed (Fig. 3.11). For Ile, one-bond correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$ and $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma_1$, two-bond correlations of $^{13}\text{C}\beta$ - $^{13}\text{C}\delta_1$ and $^{13}\text{C}\beta$ - $^{13}\text{C}'$, three-bond correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta_1$ and $^{13}\text{C}\gamma_1$ - $^{13}\text{C}'$, and four-bond

correlation of $^{13}\text{C}\delta 1$ - $^{13}\text{C}'$ are expected (Fig. 3.8). Consistently, the correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$, $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma 1$, and $^{13}\text{C}\beta$ - $^{13}\text{C}'$ were observed for Ile88 (Fig. 3.11). However, the $^{13}\text{C}\beta$ - $^{13}\text{C}\delta 1$ correlation was too weak to be visible (Fig. 3.11), probably due to the conformation dynamics of the $^{13}\text{C}\delta 1$ of Ile88. In fact, it was found that Ile88 $^{13}\text{C}\delta 1$ has double resonances (black spectrum in Fig. 3.2). In addition, all the three-bond and four-bond correlations were too weak to be visible. For Lys, a) two-bond correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$, $^{13}\text{C}\gamma$ - $^{13}\text{C}\epsilon$, and $^{13}\text{C}\beta$ - $^{13}\text{C}'$, b) three-bond correlations of $^{13}\text{C}\gamma$ - $^{13}\text{C}'$, $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta$, and $^{13}\text{C}\beta$ - $^{13}\text{C}\epsilon$, c) four-bond correlation of $^{13}\text{C}\alpha$ - $^{13}\text{C}\epsilon$, and d) five-bond correlation of $^{13}\text{C}\epsilon$ - $^{13}\text{C}'$ are expected (Fig. 3.9). In fact, only the correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ and $^{13}\text{C}\gamma$ - $^{13}\text{C}\epsilon$ were observed for Lys58, while the other correlations were too weak to be observable (Fig. 3.11). Taking Phe as an example for Phe and Tyr, two-bond correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ and $^{13}\text{C}\gamma$ - $^{13}\text{C}\epsilon 1/2$, and four-bond correlation of $^{13}\text{C}\alpha$ - $^{13}\text{C}\epsilon 1/2$ are expected (Fig. 3.7). In fact, only the $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ correlation was observed, while the $^{13}\text{C}\alpha$ - $^{13}\text{C}\epsilon 1/2$ correlation was too weak to be visible (Fig. 3.11). Taking Gln as an example for AKG derived amino acids, two-bond correlations of $^{13}\text{C}\beta$ - $^{13}\text{C}\delta$ and $^{13}\text{C}\beta$ - $^{13}\text{C}'$, three-bond correlation of $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta$, and four-bond correlation of $^{13}\text{C}\delta$ - $^{13}\text{C}'$ are expected (Fig. 3.8). Indeed, the correlations of $^{13}\text{C}\beta$ - $^{13}\text{C}\delta$ and $^{13}\text{C}\beta$ - $^{13}\text{C}'$ were observed, while $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta$ and $^{13}\text{C}\delta$ - $^{13}\text{C}'$ were too weak to be visible for Gln62 (Fig. 3.11). For Val, one-bond correlation $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$ is expected (Fig. 3.7). Surprisingly, the correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$ and $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma 2$ were observed for Val48 (Fig. 3.11). The presence of the $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma 2$ correlation is due to the scrambling (see Part II for more details). For the other residues, i.e., Ala, Ser, Cys, and Gly, no intra-residue correlation is expected (Fig. 3.7). Cys are not in the primary sequence of $m\alpha S$ (Fig. 1.4). Indeed, no cross peak was observed for Ala, Ser, and Gly, instead, only $\text{C}\alpha$ is ^{13}C labeled as indicated by the strong diagonal peaks (Fig. 3.11). Taken together, the resulting 2D ^{13}C - ^{13}C PDSO ssNMR spectrum with a short mixing time of 100 ms (Fig. 3.11) exhibits the expected ^{13}C - ^{13}C correlations generally. Considering that the cross peak intensity is proportional to the inverse cube of the inter-nuclear distance involved (i.e., r^{-3}) and the enrichment level, it is reasonable that some of the expected correlations were too weak to be visible with the short mixing time used in Fig. 3.11.

In addition to the rapid identification of residue spin system according to their expected characteristic correlations, the sparse ^{13}C labeling scheme based on $[2\text{-}^{13}\text{C}]\text{Glc}$ also facilitates sequential transfer, e.g., $^{13}\text{C}\alpha(i)$ - $^{13}\text{C}\alpha(i\pm 1)$, by the diluted ^{13}C spins, which

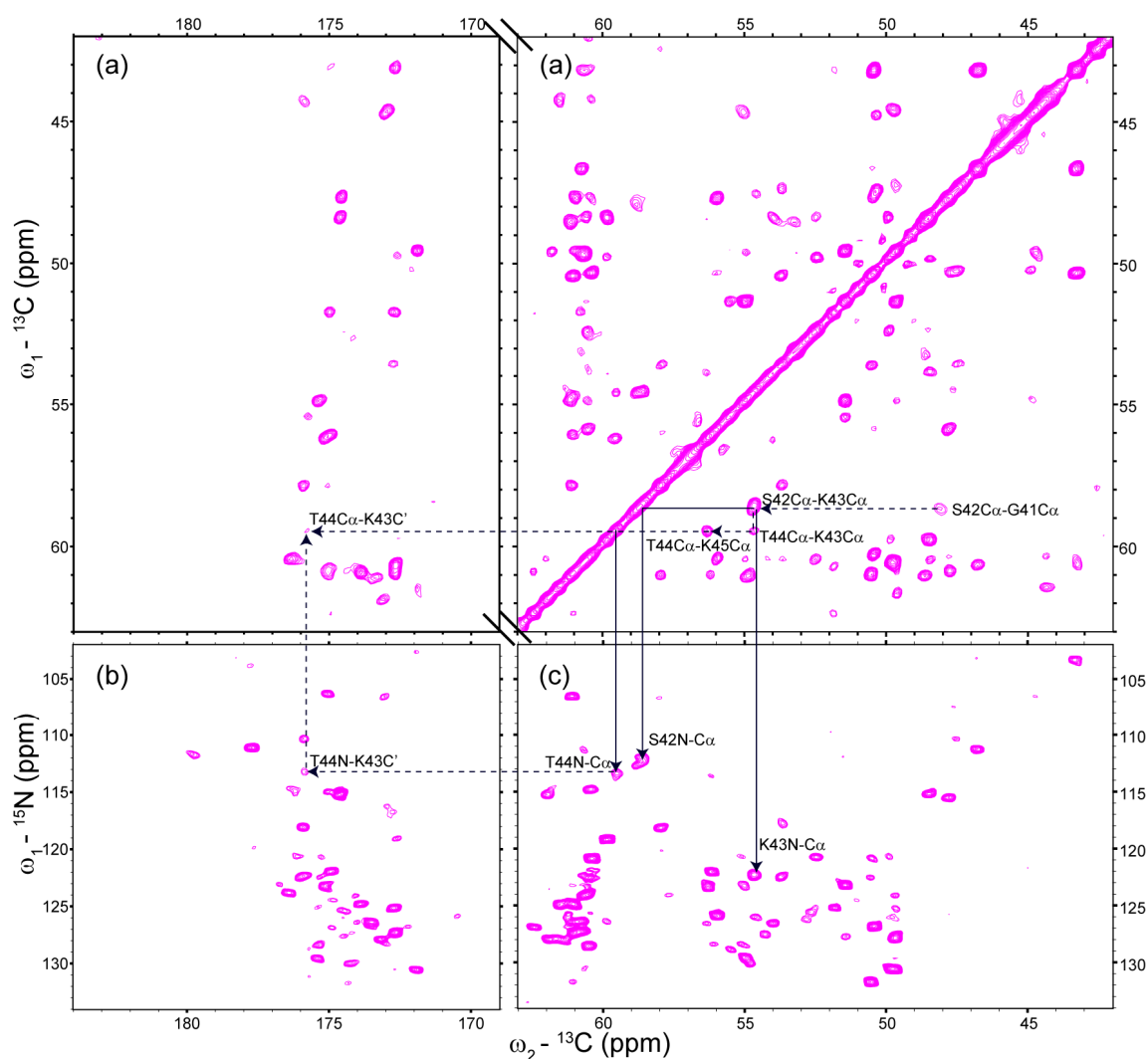


Figure 3.12. Sequential assignment strategy for backbone resonances using $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils. (a) Carbonyl and aliphatic regions of a 2D ^{13}C - ^{13}C PDSD spectrum recorded with a mixing time of 500 ms, (b) 2D NCO, and (c) 2D NCA spectra. To exemplify the assignment strategy, the sequential backbone assignment of Gly41-Thr44 is shown. Intra- and inter-residue connectivities are indicated with solid and dashed arrows, respectively. All the spectra were recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. The 2D ^{13}C - ^{13}C PDSD spectrum was zero-filled to 4096 (direct dimension) and 2048 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions. The 2D NCA and NCO spectra were zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

further permits unambiguous sequential resonance assignment in a straightforward way. As illustrated in Fig. 3.12, we propose a strategy for the sequential backbone resonance assignment based on 2D ^{13}C - ^{13}C and ^{13}C - ^{15}N correlation spectra recorded on $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils. The observed linewidths for ^{13}C and ^{15}N are ~ 0.3 ppm and ~ 0.5 ppm, respectively. The aliphatic and carbonyl regions of a 2D ^{13}C - ^{13}C PDSD spectrum with a long mixing time of 500 ms are dominated by $^{13}\text{C}\alpha(i)\text{-}^{13}\text{C}\alpha(i\pm 1)$ and $^{13}\text{C}\alpha(i)\text{-}^{13}\text{C}'(i-1)$

correlations, respectively (Fig. 3.12a). Sequential assignments involve $^{13}\text{C}\alpha(i)$ - $^{13}\text{C}\alpha(i\pm 1)$ correlations can be easily connected to $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i)$ correlations in the NCA spectrum (Fig. 3.12c). Then the $^{13}\text{C}'(i-1)$ resonances can be identified via the $^{13}\text{C}\alpha(i)$ resonances in the carbonyl region of the 2D ^{13}C - ^{13}C PDS spectrum (Fig. 3.12a) and/or from the $^{15}\text{N}(i)$ resonances in the NCO spectrum (Fig. 3.12b). As an example, the dashed arrows show the sequential connections from Gly41 to Thr44, and the solid arrows show the intra-residue connections for residues Ser42 to Thr44 (Fig. 3.12).

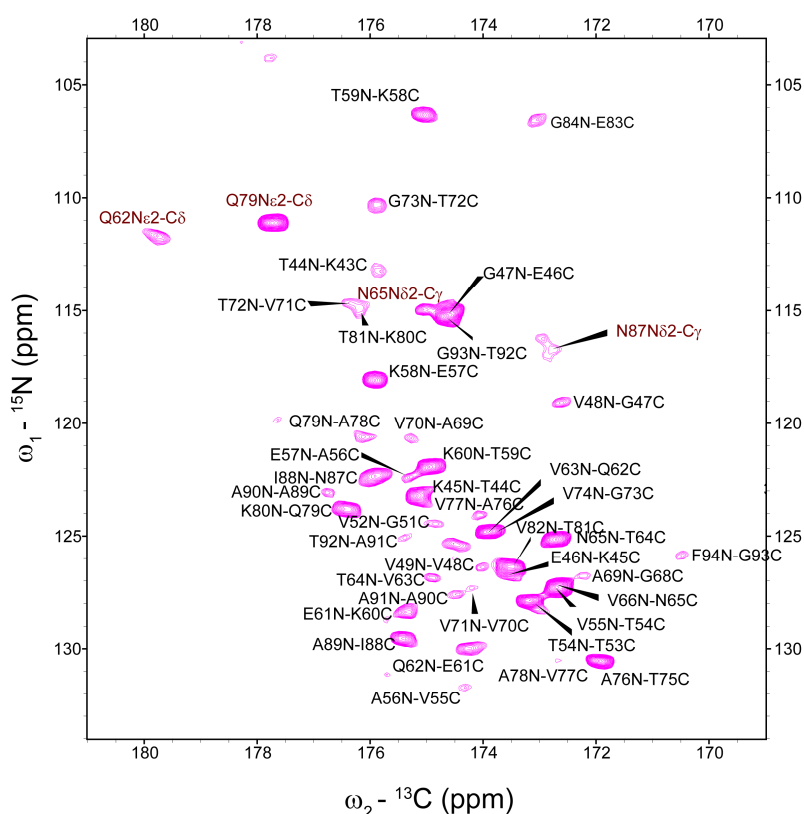


Figure 3.13. 2D NCO spectrum of $[2-^{13}\text{C}]$ Glc-labeled $m\alpha\text{S}$ fibrils. $^{15}\text{N}(i)$ - $^{13}\text{C}'(i-1)$ correlations are indicated in black, while the correlations of ^{15}N to its sidechain $^{13}\text{C}'$, i.e., $^{15}\text{N}\epsilon 2(i)$ - $^{13}\text{C}\delta(i)$ in Gln and $^{15}\text{N}\delta(i)$ - $^{13}\text{C}\gamma(i)$ in Asn, are indicated in dark red. The spectrum was recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. The spectrum was zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

It is interesting to note that most of the $^{13}\text{C}\alpha(i)$ - $^{13}\text{C}'(i-1)$ correlations observed in the carbonyl region of the 2D ^{13}C - ^{13}C PDS spectrum (Fig. 3.12a), and the strong cross peaks of $^{15}\text{N}(i)$ - $^{13}\text{C}'(i-1)$ observed in NCO spectrum (Fig. 3.12b) fit perfectly to the $^{13}\text{C}'$ of the group III residues. Consistently, as mentioned above (section 3.3.1 of Part I), the carbonyl carbon is ^{13}C labeled for the group III amino acids using the $[2-^{13}\text{C}]$ Glc labeling

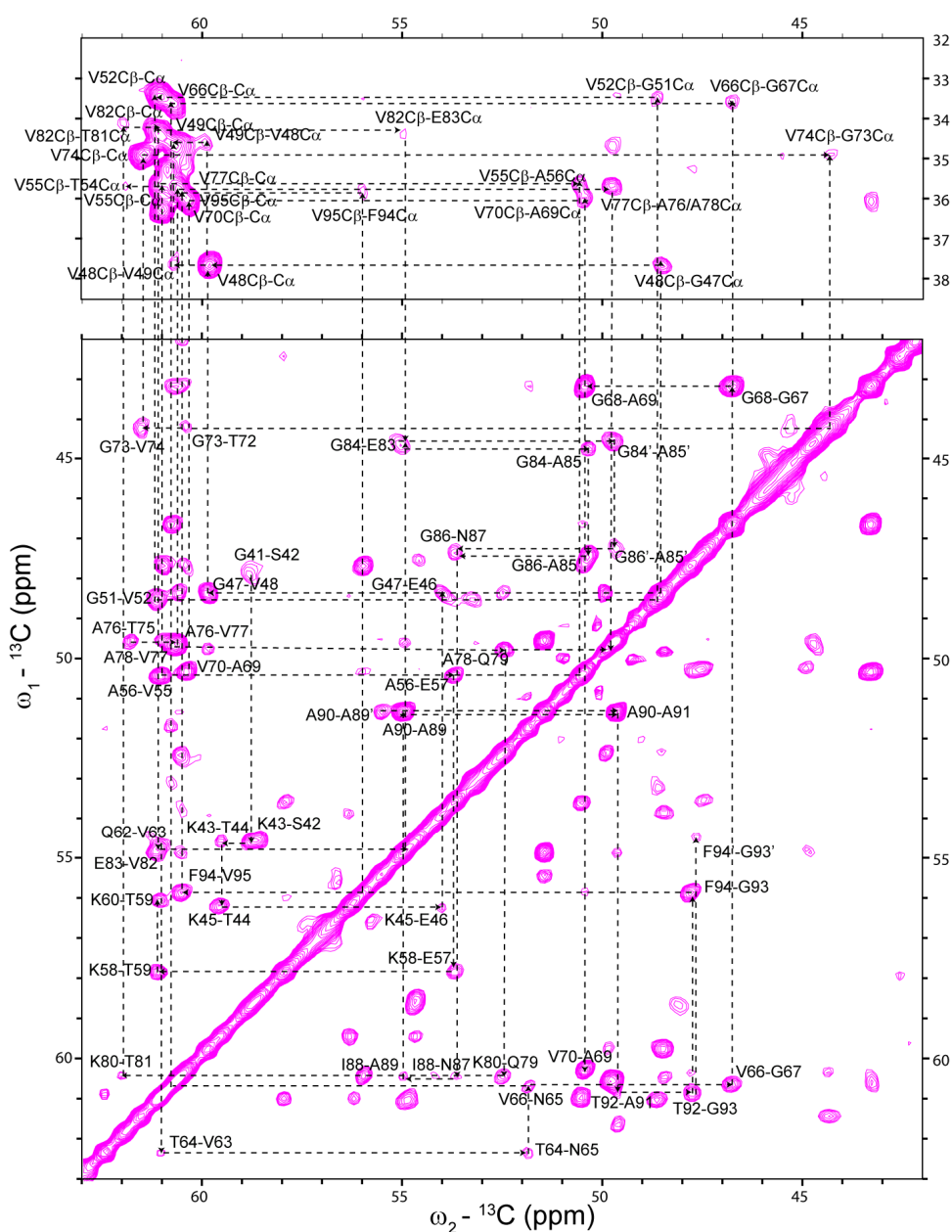


Figure 3.14. Sequential walks in a 2D ^{13}C - ^{13}C PDSO spectrum recorded with a mixing time of 500 ms of [$2\text{-}^{13}\text{C}$]Glc-labeled m α S fibrils. The sequential walks involve $^{13}\text{C}\alpha(i)$ - $^{13}\text{C}\alpha(i\pm 1)$ correlations (lower plot) and Val $^{13}\text{C}\beta(i)$ - $^{13}\text{C}\alpha(i\pm 1)$ correlations (upper plot) for the stretches Gly41-Val49, Gly51-Val52, Thr54-Lys60, Gln62-Val70, Thr72-Val74, and Thr75-Val95. The second set of sequential resonance assignments observed is indicated with primes, e.g., Phe94'. The spectra were recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. The spectrum was zero-filled to 4096 (direct dimension) and 2048 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

scheme, while without simultaneously ^{13}C labeling at the $\text{C}\alpha$ position (Fig. 3.8-3.9), which allows to observe the resonance of $^{13}\text{C}'$ with high resolution owing to the absence of one-bond ^{13}C - ^{13}C J and dipolar couplings. However, in addition to the $^{15}\text{N}(i)$ - $^{13}\text{C}'(i-1)$ correlations for the group III residues observed in the NCO spectrum (Fig. 3.13), we also

observed the $^{15}\text{N}(i)\text{-}^{13}\text{C}'(i-1)$ correlations for the group I and II residues in the NCO spectrum (Fig. 3.13), e.g., Ala, Gly and Val. From their biosynthetic pathways (Fig. A7a-b in appendix A), the $^{13}\text{C}'$ of these amino acids are derived from the carbon 1 of pyruvate or 3PG. In fact, the PPP might lead to the scrambling of $[1,2\text{-}^{13}\text{C}]$ pyruvate, $[1\text{-}^{13}\text{C}]$ pyruvate, $[1,3\text{-}^{13}\text{C}]$ pyruvate, and $[3\text{-}^{13}\text{C}]$ pyruvate (more details are presented in Part II). The $^{15}\text{N}(i)\text{-}^{13}\text{C}'(i-1)$ correlations are thus expected for the group I and II residues. Consistently, these correlations are observed with less intensity compared to that of the group III residues (Fig. 3.13), since they are only observed due to scrambling. Nevertheless, these two spectra allowed us to obtain a lot of unambiguous sequential resonance assignments regarding $^{13}\text{C}'$.

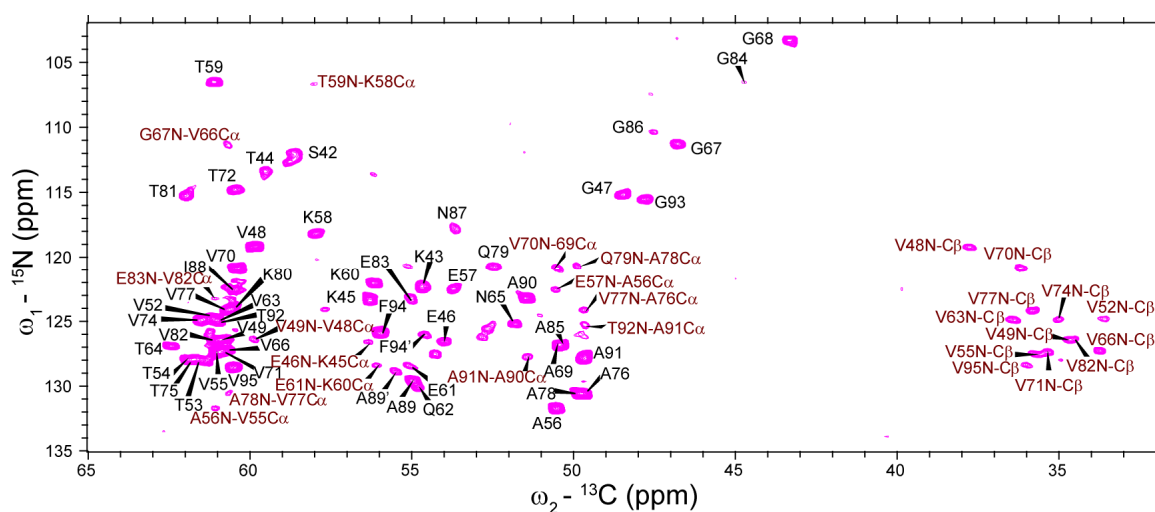


Figure 3.15. 2D NCA spectrum of $[2\text{-}^{13}\text{C}]$ Glc-labeled αS fibrils. $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)$ correlations are indicated in black, while $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i-1)$ correlations and $^{15}\text{N}(i)\text{-}^{13}\text{C}\beta(i)$ correlations of Val are indicated in dark red. The spectrum was recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. The spectrum was zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

Additionally, it is worth to note that $\text{C}\alpha$ and $\text{C}\beta$ are ^{13}C labeled simultaneously for Val using this labeling scheme, which enabled us to obtain and confirm the assignments of Val by sequential $^{13}\text{C}\beta(i)\text{-}^{13}\text{C}\alpha(i\pm 1)$ correlations (Fig. 3.14). Also, the presence of $^{15}\text{N}(i)\text{-}^{13}\text{C}\beta(i)$ correlations for Val in the NCA spectrum due to the attenuation of dipolar truncation effect [52] by the sparse ^{13}C labeling scheme allowed for the unambiguous assignments of $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)$ correlations of Val, although they heavily overlap with the ones of Thr (Fig. 3.15). The reduced dipolar truncation effect [52] also facilitates polarization transfer from ^{15}N simultaneously to preceding $^{13}\text{C}\alpha$ atoms, and helped us to

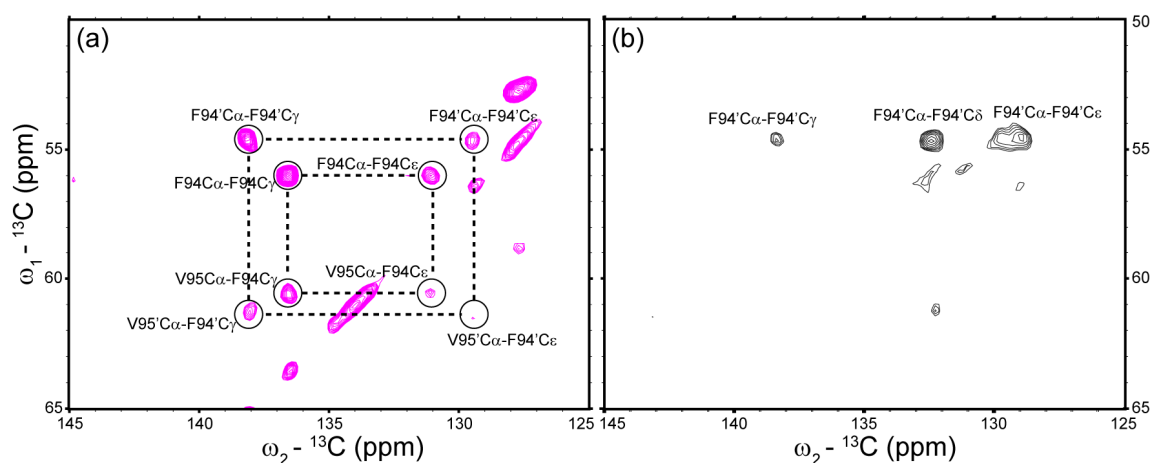


Figure 3.16. Excerpts of 2D ^{13}C - ^{13}C PDS D spectra of $m\alpha\text{S}$ fibrils, recorded on (a) a $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled sample with a mixing time of 400 ms and (b) a $[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled sample with a mixing time of 100 ms. Both spectra were recorded at 14.1 Tesla (^1H resonance frequency: 600 MHz) and at a spinning frequency of 11 kHz, but with different numbers of scans (128 scans for (a), and 192 scans for (b)). Both spectra were zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

obtain and confirm sequential resonance assignments regarding $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i-1)$ (Fig. 3.15, in red).

Moreover, significant improvement was observed for the aromatic region in the 2D ^{13}C - ^{13}C PDS D spectra of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $m\alpha\text{S}$ fibrils. The transfer efficiency from $^{13}\text{C}\alpha$ to both intra-residue and inter-residue aromatic ^{13}C atoms is enhanced by the reduced dipolar truncation effect [52], owing to the ultra-sparse ^{13}C labeling pattern by using $[2\text{-}^{13}\text{C}]\text{Glc}$ as the sole carbon source in the protein expression medium (Fig. 3.7). As an illustration, a 2D ^{13}C - ^{13}C PDS D spectrum with a mixing time of 400 ms of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $m\alpha\text{S}$ fibrils is shown in Fig. 3.16. With this long mixing time, not only the intra-residue $^{13}\text{C}\alpha(i)\text{-}^{13}\text{C}\gamma(i)$ and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\epsilon_{1/2}(i)$ correlations for Phe94, but also the sequential $^{13}\text{C}\alpha(i+1)\text{-}^{13}\text{C}\gamma(i)$ and $^{13}\text{C}\alpha(i+1)\text{-}^{13}\text{C}\epsilon_{1/2}(i)$ correlations for Phe94 were observed (Fig. 3.16). We need to note that $^{13}\text{C}\epsilon_1$ and $^{13}\text{C}\epsilon_2$ in Phe and Tyr are magnetically equivalent due to rapid ring flips, the chemical shifts between the two spins are indistinguishable. In addition, the unassigned strong correlation observed in Fig. 3.15 might belong to His50 $^{13}\text{C}\alpha\text{-}^{13}\text{C}\delta$. However, the unambiguous assignment of His Is not obtained due to a lack of unambiguous sequential correlations. The absence of the other aromatic residues in the aromatic region of the 2D ^{13}C - ^{13}C PDS D spectrum (Fig. 3.16), e.g., Phe4, Tyr39, and Tyr107, indicates that they are too flexible to be detected.

In addition, 2D ^{15}N - ^{15}N PDSO correlation spectra with a mixing time of 6 s were conducted on $[\text{U-}^{13}\text{C}]\text{Glc-}$ (Fig. 3.17a) and $[\text{2-}^{13}\text{C}]\text{Glc-}$ labeled αS fibrils (Fig. 3.17b). The observed linewidths of ^{15}N for the spectra in Fig. 3.17a and Fig. 3.17b are ~ 0.9 ppm and ~ 0.8 ppm, respectively. Considering the longer acquisition time and the higher field strength used for the 2D ^{15}N - ^{15}N PDSO spectrum recorded on $[\text{2-}^{13}\text{C}]\text{Glc-}$ labeled αS fibrils compared to that of $[\text{U-}^{13}\text{C}]\text{Glc-}$ labeled αS fibrils (Fig. 3.17), the better resolution is reasonable. In other words, it might indicate that no significant resolution enhancement in the ^{15}N dimension was observed for the $[\text{2-}^{13}\text{C}]\text{Glc}$ labeling scheme compared to that of the $[\text{U-}^{13}\text{C}]\text{Glc}$ labeling scheme.

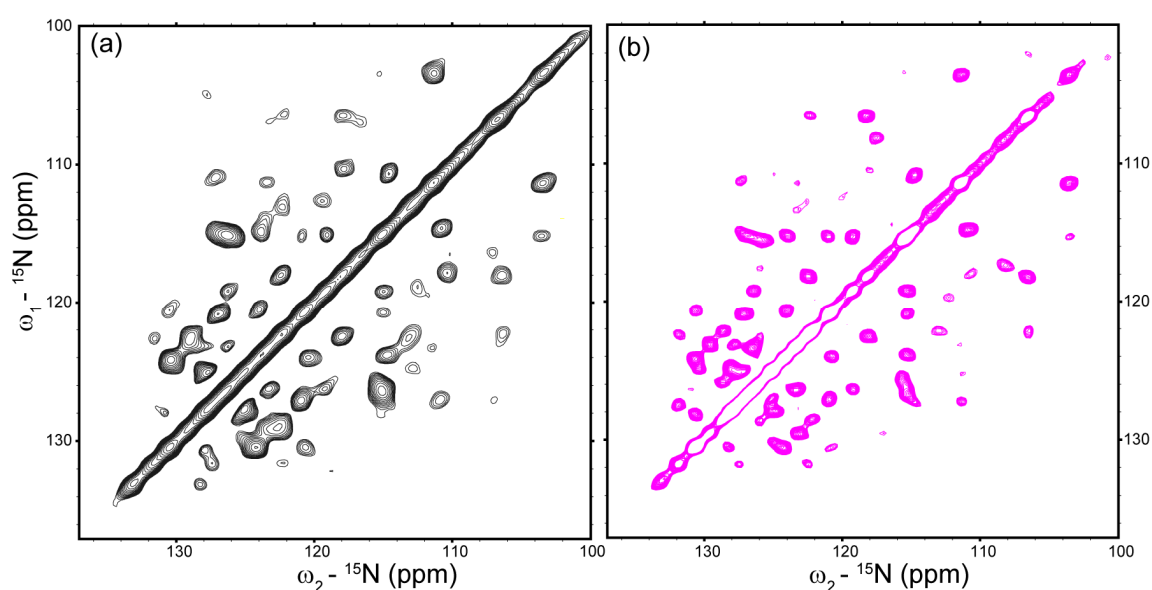


Figure 3.17. 2D ^{15}N - ^{15}N PDSO spectra with a mixing time of 6 s on (a) $[\text{U-}^{13}\text{C}]\text{Glc-}$ (in black) and (b) $[\text{2-}^{13}\text{C}]\text{Glc-}$ labeled αS fibrils (in magenta). Both spectra were recorded at a spinning frequency of 11 kHz. The black spectrum was recorded at 18.8 Tesla (^1H resonance frequency: 800 MHz) and with a maximum acquisition time of 11.5 ms (indirect dimension), while the magenta spectrum was recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz) and with a maximum acquisition time of 14.9 ms (indirect dimension). Both spectra were recorded at a spinning frequency of 11 kHz. Both spectra were zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

Taken together, the high quality of the spectra recorded on $[\text{2-}^{13}\text{C}]\text{Glc-}$ labeled αS fibrils, such as resolution enhancement, spectral simplification, and improved intra-residue and inter-residue transfer efficiencies, allows for an unambiguous sequential resonance assignment of αS fibrils. The sequential walks involving $^{13}\text{C}\alpha(i)$ - $^{13}\text{C}\alpha(i\pm 1)$ correlations and/or $^{13}\text{C}\beta(i)$ - $^{13}\text{C}\alpha(i\pm 1)$ (for Val) correlations regarding the stretches Gly41-Val49, Gly51-Val52, Thr54-Lys60, Gln62-Val70, Thr72-Val74, and Thr75-Val95 are

obtained without ambiguity and shown in Fig. 3.14. The correlations of $^{15}\text{N}(i)$ - $^{13}\text{C}\alpha(i)$ and $^{15}\text{N}(i)$ - $^{13}\text{C}'(i-1)$ are assigned unambiguously in the spectra of NCA (Fig. 3.15) and NCO (Fig. 3.13), respectively. In addition, doubling of resonances for the residues Gly84, Ala85, Gly86, Ala89, Gly93, and Phe94 was revealed by the sufficient resolution resulting from the sparse glucose labeling scheme (Fig. 3.14). A single set of resonances is observed for the other residues, indicating that no polymorphism exists in our samples, whereas the presence of the few doubled resonances may suggest local disorder of these residues. This correlates quite well with their location in the secondary structure of αS fibrils: Gly84 and Ala85 are close to one loop that includes Gly86 and Ala89, while Gly93 and Phe94 are located at the end of the structured region of αS fibrils (see the discussion about the secondary structure analysis in section 3.4 of Part I).

3.3.3 Labeling pattern of [1- ^{13}C]glucose labeling scheme

Similarly to the [2- ^{13}C]Glc labeling scheme, the labeling patterns of the 20 standard amino acids expressed in a minimal medium containing [1- ^{13}C]Glc as the sole carbon source, are analyzed in this section. The major metabolic pathways and individual biosynthetic pathways involved in the amino acid biosynthesis are illustrated in Fig. A9-11 (appendix A). Following the steps shown in Fig. A9a, approximately 1/2 [3- ^{13}C]pyruvate and 1/2 [^{12}C]pyruvate are expected when starting from [1- ^{13}C]Glc. The same pattern is expected for 3PG, i.e., approximately 1/2 [3- ^{13}C]3PG and 1/2 [^{12}C]3PG are generated during glycolysis (Fig. A9a). Consequently, about 1/2 $\text{C}\beta$ is ^{13}C enriched while no ^{13}C is labeled at the $\text{C}\alpha$ position simultaneously for the group I amino acids except Val and Leu (Fig. 3.18a). For Val, the isotopomers of $^{13}\text{C}\gamma_1$ - $^{13}\text{C}\gamma_2$ (1/4), $^{13}\text{C}\gamma_1$ - $^{12}\text{C}\gamma_2$ (1/4), $^{12}\text{C}\gamma_1$ - $^{13}\text{C}\gamma_2$ (1/4), and $^{12}\text{C}\gamma_1$ - $^{12}\text{C}\gamma_2$ (1/4) are formed. For Leu, the isotopomers of $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta_1$ - $^{13}\text{C}\delta_2$ (1/8), $^{13}\text{C}\alpha$ - $^{13}\text{C}\delta_1$ - $^{12}\text{C}\delta_2$ (1/8), $^{13}\text{C}\alpha$ - $^{12}\text{C}\delta_1$ - $^{13}\text{C}\delta_2$ (1/8), $^{12}\text{C}\alpha$ - $^{13}\text{C}\delta_1$ - $^{13}\text{C}\delta_2$ (1/8), $^{13}\text{C}\alpha$ - $^{12}\text{C}\delta_1$ - $^{12}\text{C}\delta_2$ (1/8), $^{12}\text{C}\alpha$ - $^{12}\text{C}\delta_1$ - $^{13}\text{C}\delta_2$ (1/8), $^{12}\text{C}\alpha$ - $^{13}\text{C}\delta_1$ - $^{12}\text{C}\delta_2$ (1/8), and $^{12}\text{C}\alpha$ - $^{12}\text{C}\delta_1$ - $^{12}\text{C}\delta_2$ (1/8) are formed (Fig. 3.18a).

As mentioned in section 3.3.1 of Part I, the carbon skeletons of the group II amino acids stem from PEP, 3PG, E4P and R5P. PEP and 3PG, the intermediates derived from glycolysis, are ^{13}C enriched at carbon 3 with an enrichment level of about 1/2 when starting from [1- ^{13}C]Glc (Fig. A9a). An enrichment level of about 1/2 for $^{13}\text{C}\epsilon_1$ of His,

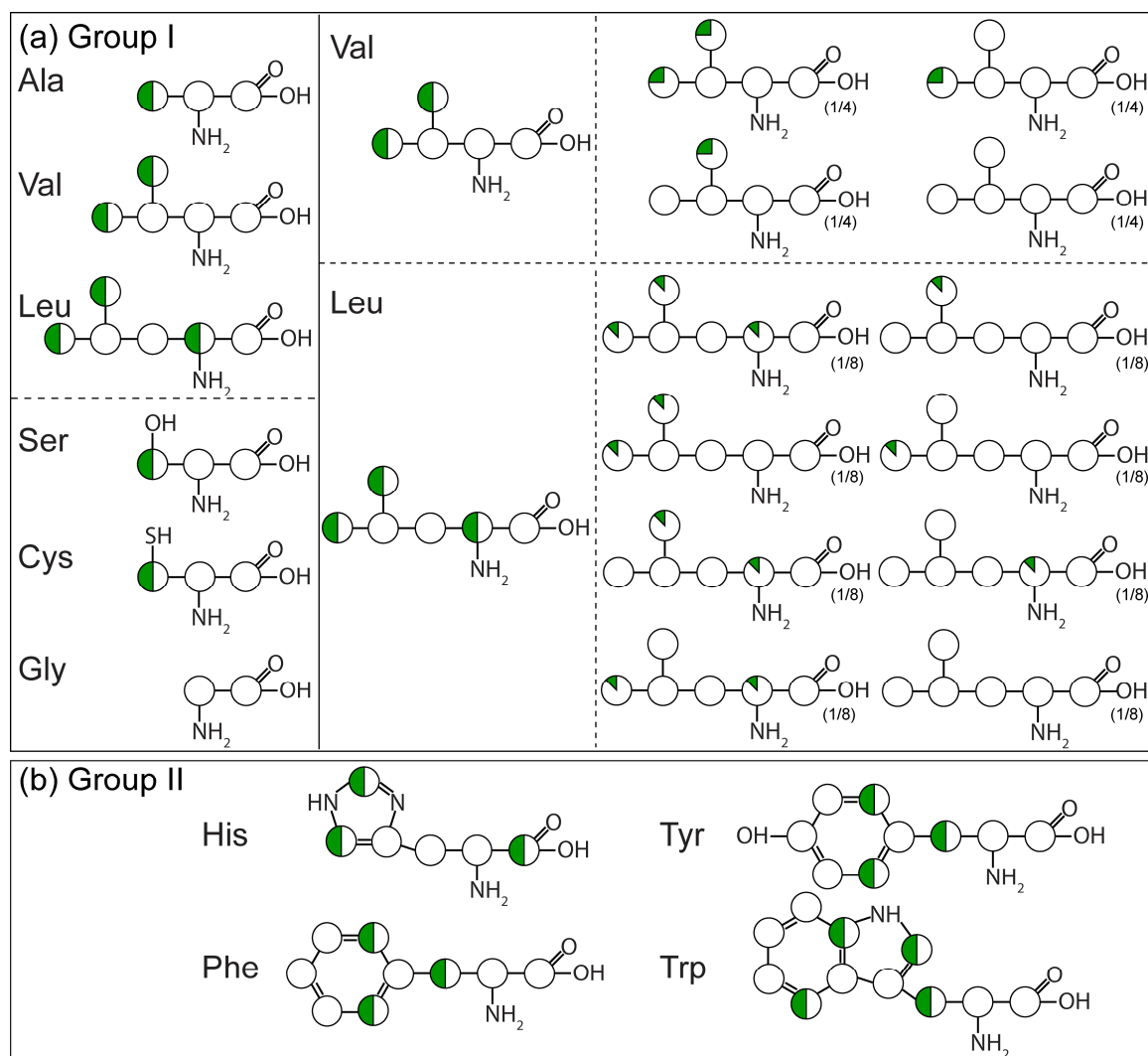


Figure 3.18. Schematic representation of the labeling patterns for amino acids synthesized from (a) the glycolytic intermediates (i.e., group I amino acids) and (b) the PPP intermediates as well as the glycolytic intermediates (i.e., group II amino acids) in a $[1-^{13}\text{C}]\text{Glc}$ containing medium. Carbons that are ^{13}C labeled are filled with green color, and the enrichment level is indicated by the percentage of the filled color. The percentage of the ^{13}C spins derived from E4P or R5P, i.e., His $^{13}\text{C}\delta 2/^{13}\text{C}'$, Phe/Tyr $^{13}\text{C}\delta 2$ (or $^{13}\text{C}\delta 1$), and Trp $^{13}\text{C}\delta 1/^{13}\text{C}\epsilon 2$, could not be determined, a 1/2 distribution is instead displayed, and their isotopomers are not considered here. Val and Leu are the exceptions of group I amino acids, and the detailed isotopomers are displayed in right. For clarity, the percentage of each isotopomer of Val and Leu is shown below.

$^{13}\text{C}\beta$ and $^{13}\text{C}\delta 1$ (or $^{13}\text{C}\delta 2$) of Phe and Tyr, and $^{13}\text{C}\beta$ and $^{13}\text{C}\epsilon 3$ of Trp is thus expected, since they are derived from PEP and 3PG (Fig. A10c). E4P and R5P are generated in the PPP, and will be ^{13}C enriched at position 4 and positions 1 and 5 when starting from $[1-^{13}\text{C}]\text{Glc}$, respectively (Fig. A9b). Thereby, a) $^{13}\text{C}'$ - $^{13}\text{C}\delta 2$ spin pair of His are expected from $[1,5-^{13}\text{C}]\text{R5P}$, b) $^{13}\text{C}\delta 2$ (or $^{13}\text{C}\delta 1$) of Phe and Tyr are expected from $[4-^{13}\text{C}]\text{E4P}$, and c) $^{13}\text{C}\delta 1$ and $^{13}\text{C}\epsilon 2$ of Trp are expected from $[1,5-^{13}\text{C}]\text{R5P}$ and $[4-^{13}\text{C}]\text{E4P}$ (Fig. A10c). In summary, the expected labeling patterns, including the isotopomers and the relative

incorporation levels of individual isotopomers for the group I and II amino acids expressed in a $[1-^{13}\text{C}]$ Glc containing medium are illustrated in Fig. 3.18.

Likewise, the expected labeling patterns of the group III amino acids (i.e., Asp, Asn, Thr, Ile, Met, Lys, Glu, Gln, Pro, and Arg) grown in a $[1-^{13}\text{C}]$ Glc containing medium for protein expression, are illustrated in Fig. 3.19-20. As shown in Fig. A11b, 6 types of OA (with one ^{13}C label or more ^{13}C labels) are yielded during TCA cycle when using $[1-^{13}\text{C}]$ Glc as the sole carbon source in the protein expression medium, $[3-^{13}\text{C}]$ OA, $[1,3-^{13}\text{C}]$ OA, $[2,4-^{13}\text{C}]$ OA, $[2,3-^{13}\text{C}]$ OA, $[1,2,3-^{13}\text{C}]$ OA, and $[2,3,4-^{13}\text{C}]$ OA. As

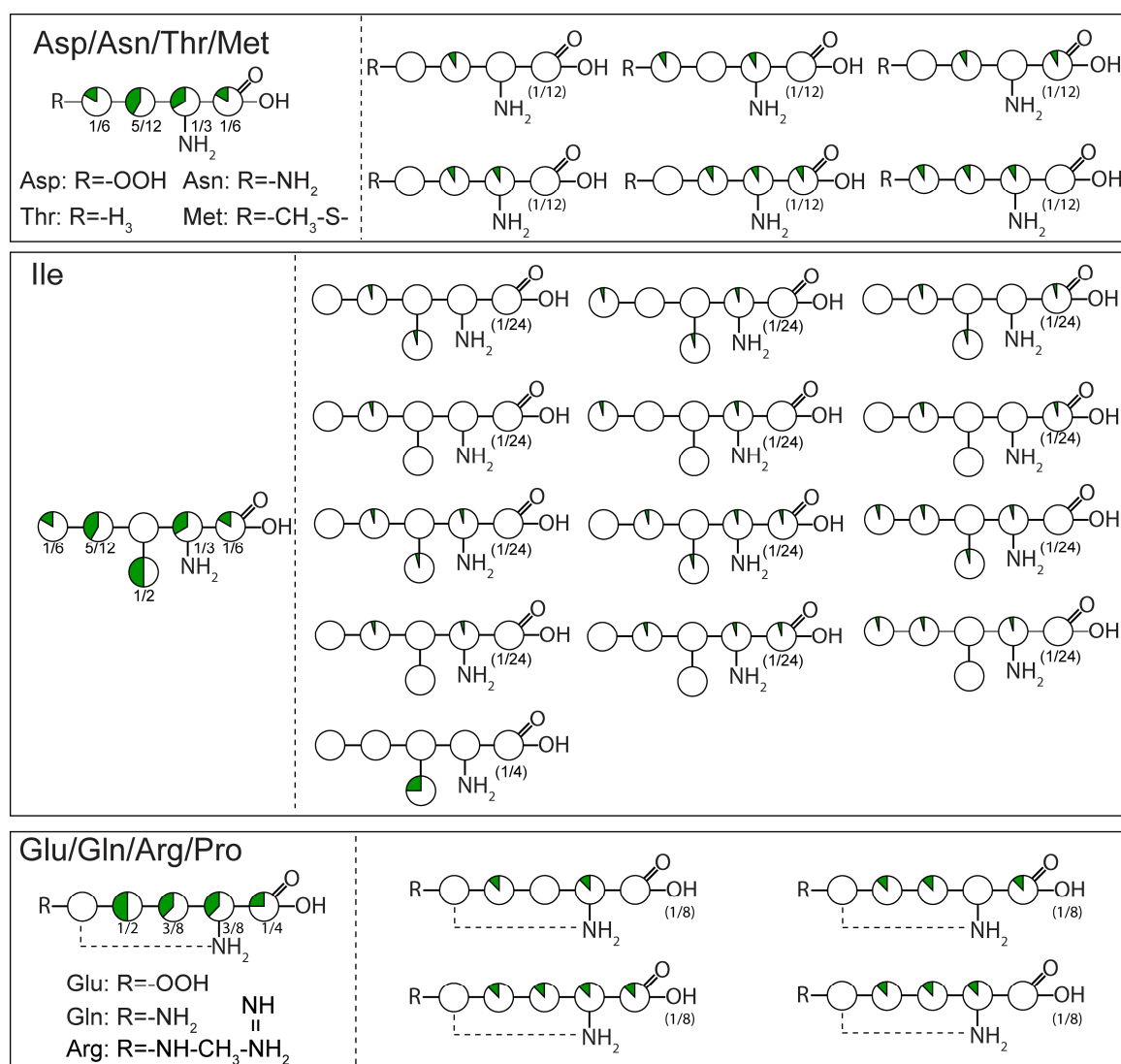


Figure 3.19. Schematic representation of the expected labeling pattern for group III amino acids produced in a $[1-^{13}\text{C}]$ Glc containing medium for protein expression. Carbons that are ^{13}C labeled are filled with green color, and the enrichment level is indicated by the percentage of the filled color. For clarity, only the isotopomers with at least one ^{13}C label are shown. The list of isotopomers for Lys is shown in Fig. 3.20.

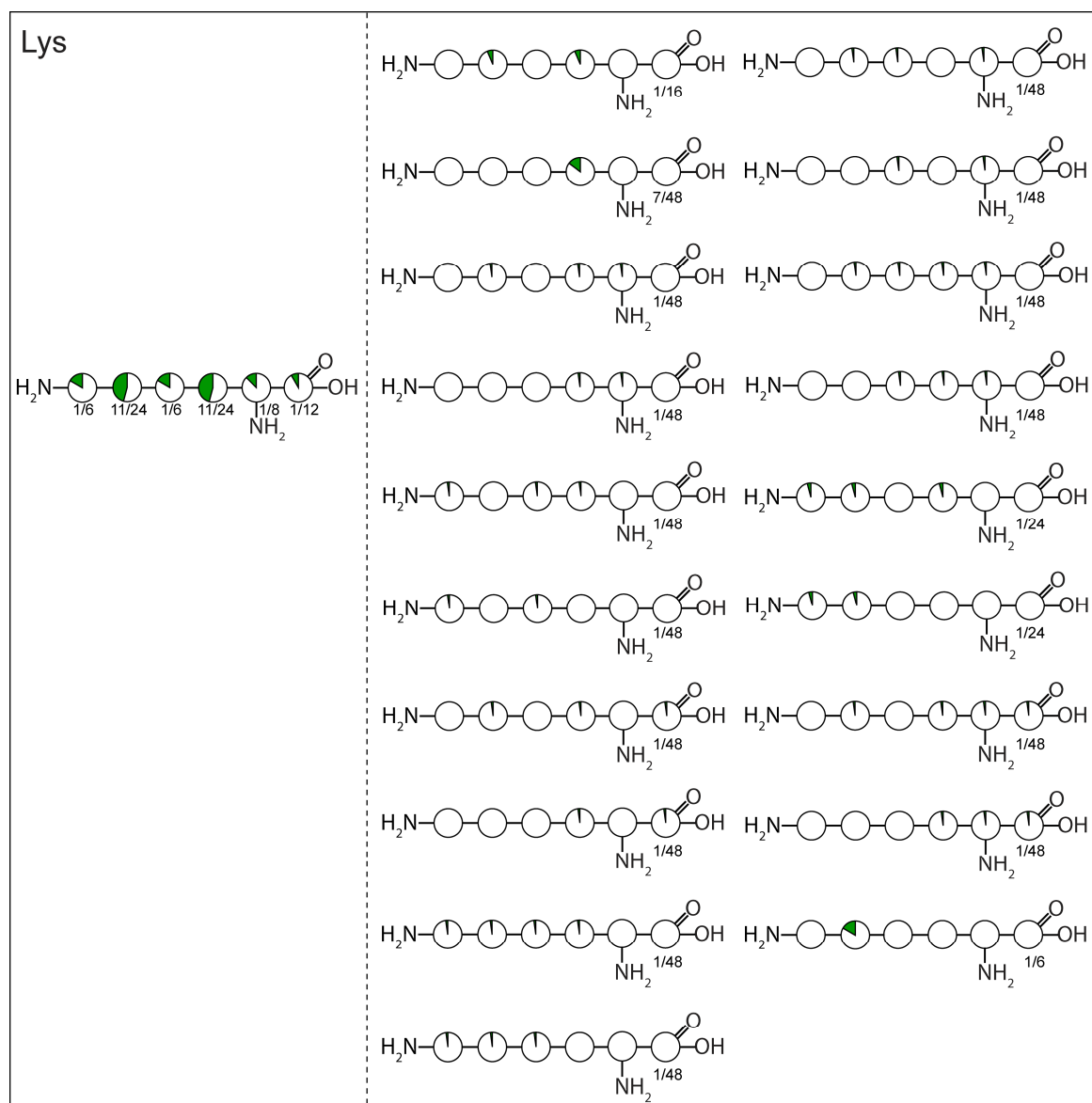


Figure 3.20. The isotopomers of Lys produced in a $[1-^{13}\text{C}]\text{Glc}$ containing medium for protein expression. Carbons that are ^{13}C labeled are filled with green color, and the enrichment level is indicated by the percentage of the filled color. For clarity, only the isotopomers with at least one ^{13}C label are shown.

a result, 6 isotopomers with at least one ^{13}C spin, i.e., $^{13}\text{C}\beta$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma$, $^{13}\text{C}\beta\text{-}^{13}\text{C}'$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma$, and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta\text{-}^{13}\text{C}'$ are thus formed for OA derived amino acids with the exceptions of Ile and Lys. The enrichment level of each isotopomer (with at least one ^{13}C spin) is estimated to be $1/12$ (i.e., $1/6 \times 1/2$). As mentioned in section 3.3.1 of Part I; the labeling patterns for Ile and Lys are more complicated due to their biosynthetic pathways (Fig. A11a). In short, 13 isotopomers with at least one ^{13}C label are expected for Ile (Fig. 3.19), and 19 isotopomers with at least one ^{13}C label are expected for Lys (Fig. 3.20). As shown in Fig. A11d, 4 types of AKG (with ^{13}C labels) are generated in

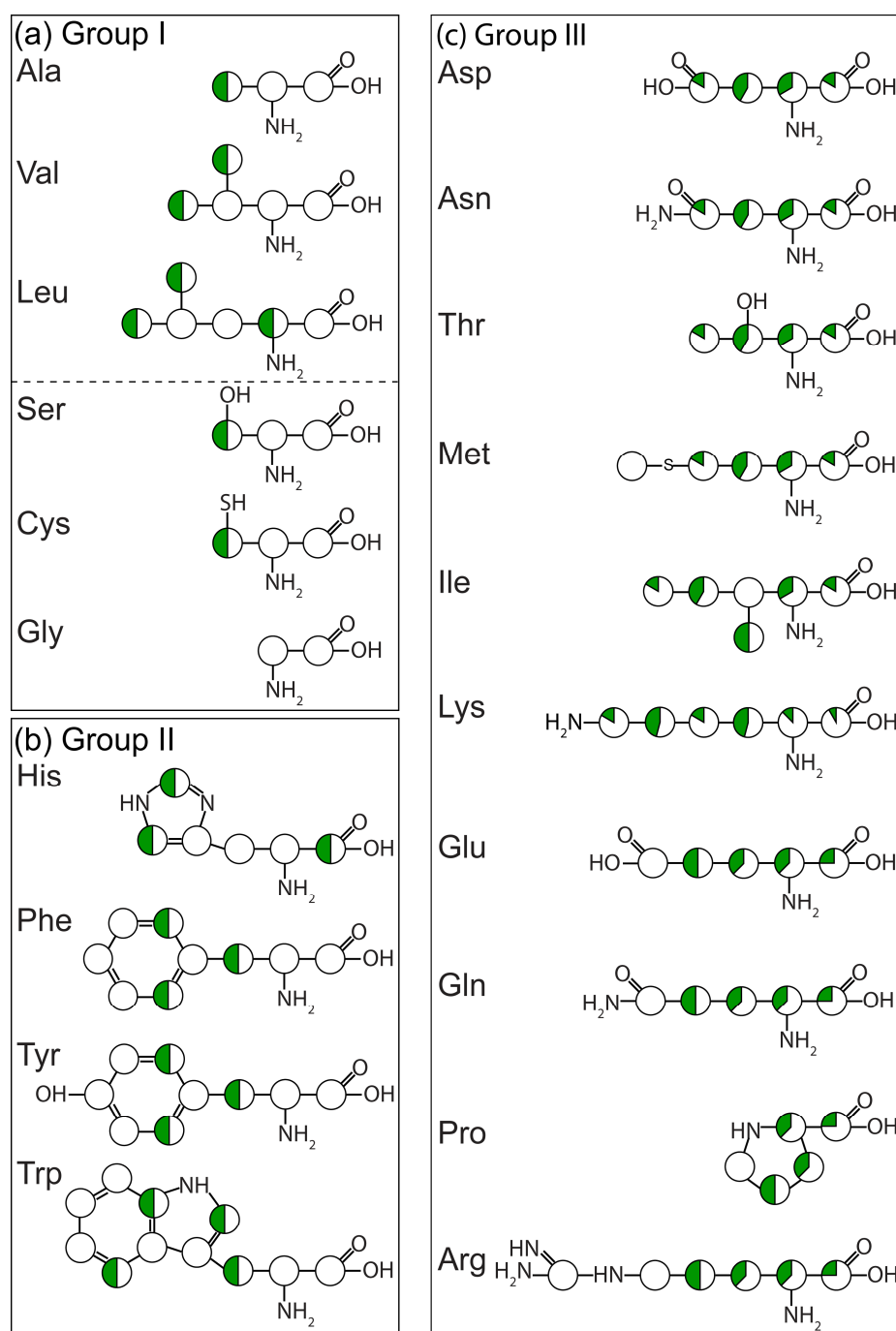


Figure 3.21. Schematic representation of the expected labeling pattern for the 20 amino acids produced in a $[1-^{13}\text{C}]\text{Glc}$ containing medium for protein expression. (a) Group I, (b) group II, and (c) group III amino acids. Carbons that are ^{13}C labeled are filled with green color, and the enrichment level is indicated by the percentage of the filled color. For the amino acids that have multiple labeling patterns, the detailed isotopomers are displayed in Fig. 3.18-3.20.

total during TCA cycle when starting from $[1-^{13}\text{C}]\text{Glc}$. Thereby, 4 isotopomers with ^{13}C isotoped spin pairs, i.e., $^{13}\text{C}\alpha-^{13}\text{C}\gamma$, $^{13}\text{C}\beta-^{13}\text{C}\gamma-^{13}\text{C}'$, $^{13}\text{C}\alpha-^{13}\text{C}\beta-^{13}\text{C}\gamma$, and $^{13}\text{C}\alpha-^{13}\text{C}\beta-^{13}\text{C}\gamma-^{13}\text{C}'$ are thus formed for Glu, Gln, Pro, and Arg. The incorporation level of each isotopomer is estimated to be about 1/8 (i.e., $1/4 \times 1/2$, Fig. 3.20).

Taken together, as shown in Fig. 3.21, a new labeling pattern sheet for all the 20 amino acids based on protein expression in an *E.coli* medium containing [1-¹³C]Glc as the sole carbon source, is proposed. Generally, the [1-¹³C]Glc labeling scheme preferentially labels the ends of the sidechains, which is in a complementary manner compared with that of the [2-¹³C]Glc labeling scheme. For group I and group II amino acids, an alternating labeling pattern is achieved, i.e., ¹²C-¹³C-¹²C or ¹³C-¹²C-¹³C. For group III amino acids, multiple ¹³C labels, e.g., ¹³C-¹³C or ¹³C-¹³C-¹³C or ¹³C-¹³C-¹³C-¹³C, are achieved, but with a much lower enrichment level (Fig. 3.19-20) compared to that of the [1,3-¹³C]glycerol labeling scheme (Fig. 1.3).

3.3.4 Resolution enhancement and spectrum simplification using [1-¹³C]glucose labeling scheme

Similarly to the [2-¹³C]Glc labeling scheme [61], the ultra-sparse labeling scheme based on [1-¹³C]Glc [58; 61] also leads to enhanced resolution of ssNMR spectra due to the removal of the majority of one-bond ¹³C-¹³C J and dipolar couplings. The spectral overlap is alleviated dramatically by a reduced number of cross peaks as well. More importantly, as a perfectly complementary labeling scheme to the [2-¹³C]Glc labeling scheme [61], the [1-¹³C]Glc labeling scheme [58; 61] enables one to detect the ¹³C spins that are not labeled by the [2-¹³C]Glc labeling scheme, e.g., ¹³C β for Ala, Ser, Cys, Phe, Tyr, and Trp, and ¹³C α for Leu. However, Cys and Trp are not in the primary sequence of m α S fibrils (Fig. 1.4). Leu and Tyr were not unambiguously assigned due to a lack of sequential connections. Moreover, both the backbone and sidechain assignments of the group III amino acids, i.e., Asn, Asp, Thr, Met, Lys, Ile, Glu, Gln, Arg, and Pro could be readily identified. Arg is not in the primary sequence of m α S fibrils (Fig. 1.4). Asp and Pro were observed but no unambiguous assignments were obtained.

As an illustration, we recorded a 2D ¹³C-¹³C PDSO spectrum with a mixing time of 100 ms. With this mixing time, the resultant ssNMR spectrum is mostly dominated by intra-residue correlations (Fig. 3.22). For the group I amino acids except Val and Leu, no intra-residue correlation should be expected (Fig. 3.18a), and indeed no visible cross peak was observed (Fig. 3.22). In contrast, Val ¹³C γ 1-¹³C γ 2 correlations were observed (Fig. 3.22). For the group II amino acids, as mentioned in section 3.3.3 of Part I, the C δ 2/C ϵ 1/C' for His, C β /C δ 1/C δ 2 for Phe and Tyr, and C β /C δ 1/C ϵ 2/C ϵ 3 for Trp are

expected to be ^{13}C labeled using the $[1-^{13}\text{C}]\text{Glc}$ labeling scheme (Fig. 3.18). Taking Phe as an example, the intra-residue $^{13}\text{C}\beta$ - $^{13}\text{C}\delta(1/2)$ correlation was observed in the 2D ^{13}C - ^{13}C PDSO spectrum with a mixing time of 100 ms (Fig. 3.22). Due to rapid ring flips of Phe and Tyr, $^{13}\text{C}\delta 1$ and $^{13}\text{C}\delta 2$ are magnetically equivalent, the chemical shifts between the two spins are thus indistinguishable.

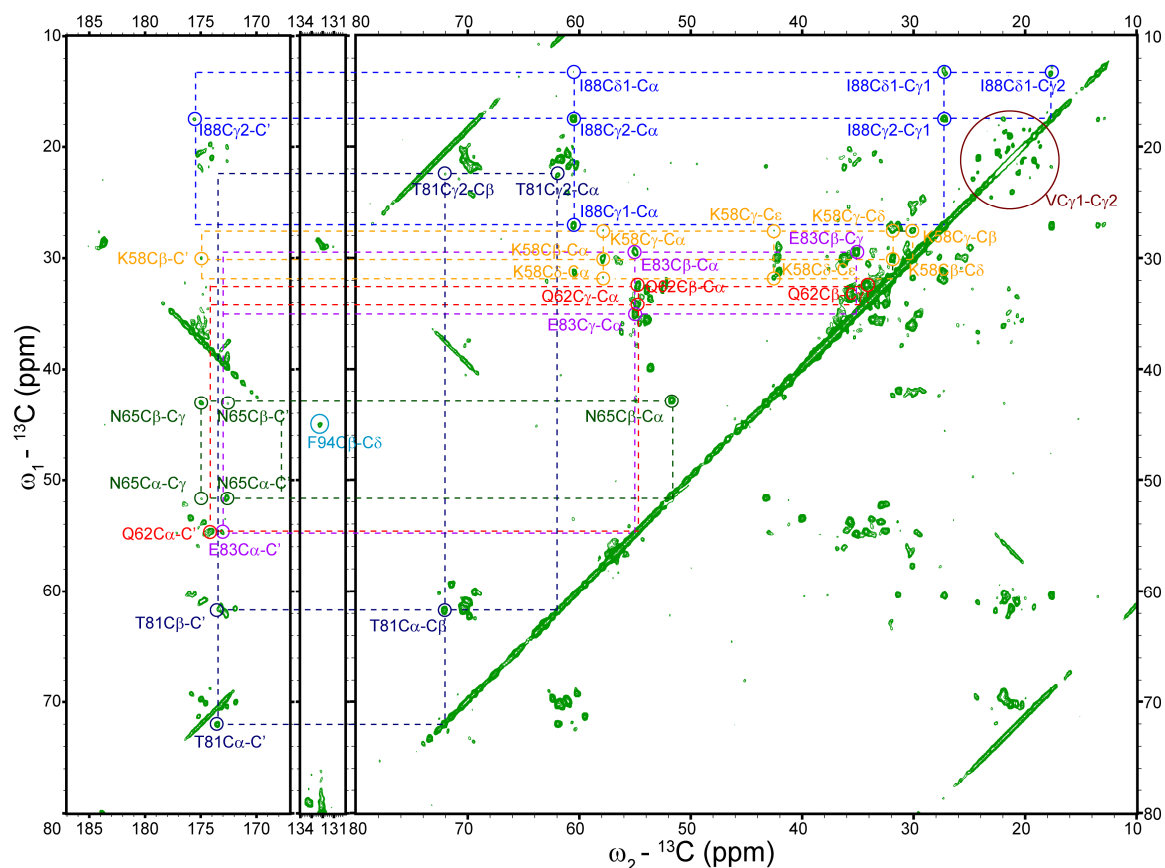


Figure 3.22. 2D ^{13}C - ^{13}C PDSO spectrum with a mixing time of 100 ms on $[1-^{13}\text{C}]\text{Glc}$ -labeled αs fibrils, conducted on a 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. The spectrum was zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions. Correlations of residue Lys58, Gln62, Asn65, Thr81, Glu83, and Ile88 for αs fibrils are illustrated with dashed lines, whereas cross peaks are indicated with circles, otherwise, no cross peak is visible or the visible cross peaks belong to other residues. Correlations of Val $^{13}\text{C}\gamma 1$ - $^{13}\text{C}\gamma 2$ are circled in dark red.

For the group III amino acids, both the backbone and sidechain assignments could be readily identified using $[1-^{13}\text{C}]\text{Glc}$ labeling scheme. As an illustration, correlations of Lys58, Gln62, Asn65, Thr81, Glu83, and Ile88 of αs fibrils are shown in Fig. 3.22. Except Ile and Lys, the same correlations are expected for OA derived amino acids, i.e., Asp, Asn, Thr, and Met. Taking Thr as an example for these amino acids, a) one-bond correlations of $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$, $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma 2$, and $^{13}\text{C}\alpha$ - $^{13}\text{C}'$, and b) two-bond correlations of

$^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}'$ are expected (Fig. 3.19). Indeed, for Thr81, the correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$, $^{13}\text{C}\beta\text{-}^{13}\text{C}'$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}'$, and $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ were observed (Fig. 3.22). For Ile, a) one-bond correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}'$ and $^{13}\text{C}\gamma_1\text{-}^{13}\text{C}\delta_1$, b) two-bond correlations of $^{13}\text{C}\gamma_1\text{-}^{13}\text{C}\gamma_2$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$, and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$, and c) three-bond correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\delta_1$, $^{13}\text{C}\gamma_1\text{-}^{13}\text{C}'$, $^{13}\text{C}\gamma_2\text{-}^{13}\text{C}\delta_1$, and $^{13}\text{C}\gamma_2\text{-}^{13}\text{C}'$ are expected (Fig. 3.19). In fact, the correlations $^{13}\text{C}\gamma_1\text{-}^{13}\text{C}\delta_1$, $^{13}\text{C}\gamma_1\text{-}^{13}\text{C}\gamma_2$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$, $^{13}\text{C}\gamma_2\text{-}^{13}\text{C}\delta_1$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\delta_1$, and $^{13}\text{C}\gamma_2\text{-}^{13}\text{C}'$ were observed, however, the $^{13}\text{C}\gamma_1\text{-}^{13}\text{C}'$ correlation was too weak to be observed (Fig. 3.22). For Lys, a) one-bond correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta$, $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma$, $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta$, $^{13}\text{C}\delta\text{-}^{13}\text{C}\epsilon$, and $^{13}\text{C}\alpha\text{-}^{13}\text{C}'$, b) two-bond correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma$, $^{13}\text{C}\gamma\text{-}^{13}\text{C}\epsilon$, $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta$, and $^{13}\text{C}\beta\text{-}^{13}\text{C}'$, and c) three-bond correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\delta$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\epsilon$, and d) four-bond correlation $^{13}\text{C}\delta\text{-}^{13}\text{C}'$ are expected (Fig. 3.20). Indeed, the $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta$, $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta$, $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma$, $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta$, $^{13}\text{C}\delta\text{-}^{13}\text{C}\epsilon$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma$, $^{13}\text{C}\beta\text{-}^{13}\text{C}'$, $^{13}\text{C}\gamma\text{-}^{13}\text{C}\epsilon$, and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\delta$ correlations were observed in Fig. 3.22. For the AKG derived amino acids (i.e., Glu, Gln, Pro and Arg), the same correlations are expected. Taking Glu as an example, a) one-bond correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta$, $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma$, and $^{13}\text{C}\alpha\text{-}^{13}\text{C}'$, b) two-bond correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}'$, c) three-bond correlation $^{13}\text{C}\gamma\text{-}^{13}\text{C}'$ are expected (Fig. 3.19). Indeed, for Glu83, the correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta$, $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma$, $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma$, and $^{13}\text{C}\alpha\text{-}^{13}\text{C}'$ were observed, while $^{13}\text{C}\beta\text{-}^{13}\text{C}'$ and $^{13}\text{C}\gamma\text{-}^{13}\text{C}'$ correlations were too weak to be observed (Fig. 3.22). Considering that the intensity is proportional to the inverse cube of the internuclear distances and the ^{13}C incorporation levels, the above mentioned observations in the spectrum of 2D $^{13}\text{C}\text{-}^{13}\text{C}$ PDSO with a mixing time of 100 ms (Fig. 3.22), are consistent with what is expected based on the metabolic pathways of *E. coli* (Fig. 3.18-20).

In addition to resolution enhancement and spectrum simplification observed in the ssNMR spectra recorded on $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils, the unambiguous sequential resonance assignment is also facilitated. Similarly to the $[2\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme, a sequential assignment strategy is proposed based on 2D $^{13}\text{C}\text{-}^{13}\text{C}$ and $^{13}\text{C}\text{-}^{15}\text{N}$ spectra of $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils (Fig. 3.23). As an example, Fig. 3.23 shows the assignment of Thr and Val based on the combination of a 2D $^{13}\text{C}\text{-}^{13}\text{C}$ PDSO spectrum with a mixing time of 700 ms recorded on $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils, and 2D $^{13}\text{C}\text{-}^{13}\text{C}$ PDSO (mixing time: 100 ms), NCO, and NCA spectra recorded on $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils. Dashed lines in Fig. 3.23 illustrate the unambiguous assignment of the

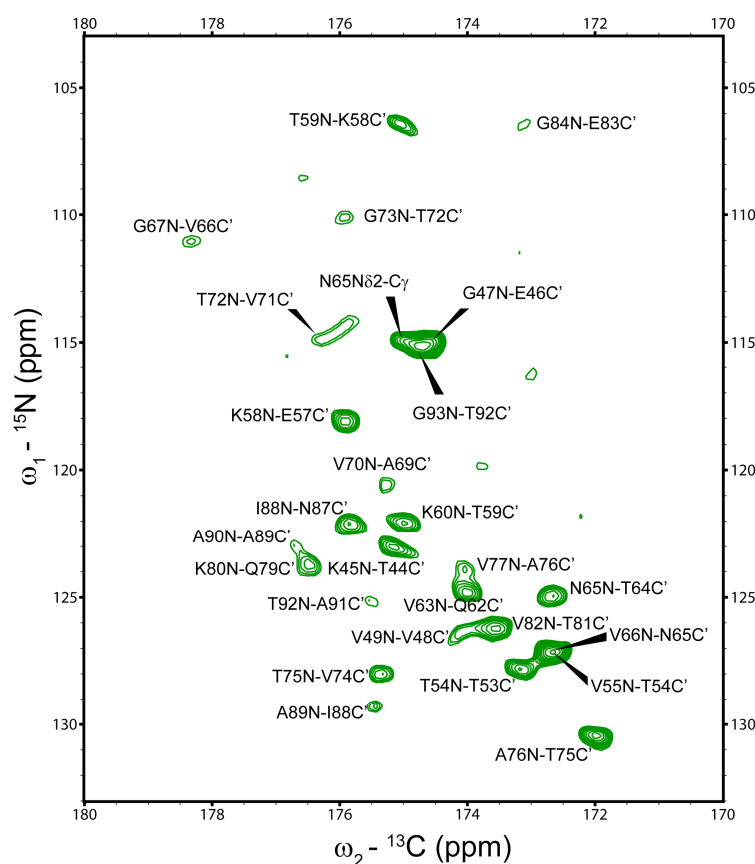


Figure 3.24. 2D NCO spectra of $[1-^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils. $^{15}\text{N}(i)-^{13}\text{C}'(i-1)$ correlations are indicated in black. The spectrum was recorded at 20.0 Tesla (^1H resonance frequency: 8500 MHz) and at a spinning frequency of 11 kHz, zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

critical residue Thr53, one of the two variant sites between $\text{h}\alpha\text{S}$ and $\text{m}\alpha\text{S}$ located in the fibril core.

It is worth to note that, the carbonyl carbons are also ^{13}C labeled for group III amino acids using the $[1-^{13}\text{C}]\text{Glc}$ labeling scheme (Fig. 3.19-20). Most carbonyl carbons are ^{13}C labeled without simultaneously being ^{13}C enriched at the $\text{C}\alpha$, e.g., isotopomers with $^{13}\text{C}\beta-^{13}\text{C}'$ and $^{13}\text{C}\beta-^{13}\text{C}\gamma_2-^{13}\text{C}'$ for OA (except Ile and Lys) and AKG derived amino acids (Fig. 3.19). However, a small amount of carbonyl carbons are ^{13}C labeled simultaneously with the $\text{C}\alpha$, e.g., the isotopomers with $^{13}\text{C}\alpha-^{13}\text{C}'$ are also formed for OA (except Ile and Lys) and AKG derived amino acids (Fig. 3.19-20). Nevertheless, the ^{13}C labeling pattern regarding the carbonyl carbons based on the $[1-^{13}\text{C}]\text{Glc}$ labeling scheme leads to the observation of the $^{13}\text{C}'$ resonances with high resolution (Fig. 3.23b-c). As an illustration, the 2D NCO spectrum recorded on $[1-^{13}\text{C}]\text{Glc}$ $\text{m}\alpha\text{S}$ fibrils is shown in Fig. 3.24. Most of the strong cross peaks were unambiguously assigned to the $^{15}\text{N}(i)-^{13}\text{C}'(i-1)$

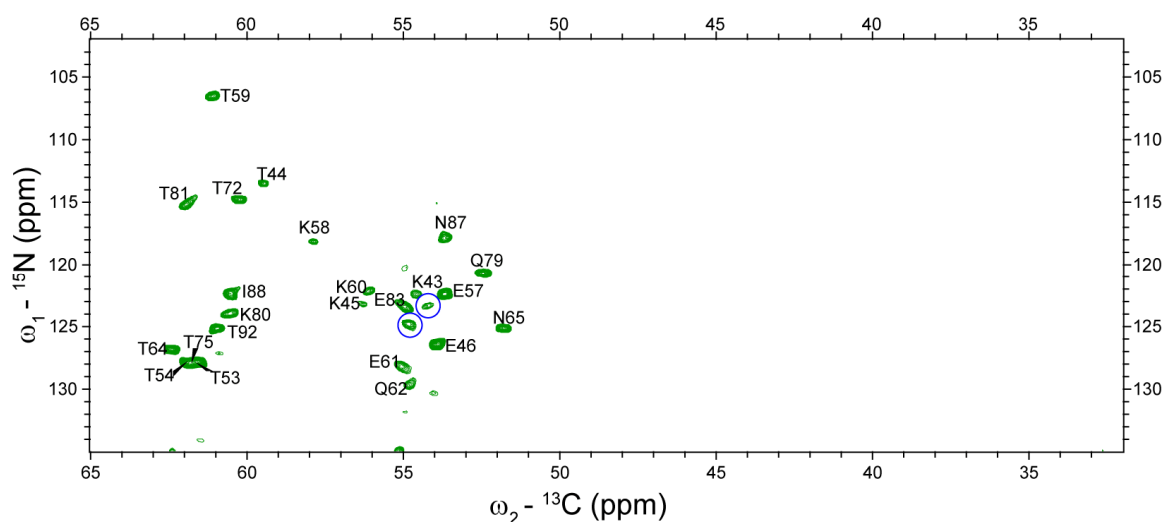


Figure 3.25. 2D NCA spectra of $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils. The spectrum was recorded at 20 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz, and zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions. The cross peaks that are only observed in the NCA spectrum of $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils, while not present in the NCA spectrum of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils (Fig. 3.14) are indicated with blue circles.

correlations of the group III amino acids. In addition, the $^{15}\text{N}(i)\text{-}^{13}\text{C}'(i-1)$ correlations for the group I and II residues, e.g., Ala, were also observed, but with much weaker intensity (Fig. 3.24).

In addition, the 2D NCA spectrum of $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils was obtained with a more reduced number of cross peaks (Fig. 3.25) compared to the NCA spectrum of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils (Fig. 3.15). Due to the fact that, the $\text{C}\alpha\text{s}$ of Ala, Ser, Gly, Cys, Phe, Tyr, and Trp are only ^{13}C enriched by the $[2\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme, and not by the $[1\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme, the comparison of the NCA spectra of $[2\text{-}^{13}\text{C}]\text{Glc}$ - (Fig. 3.15) and $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils (Fig. 3.25) thus allows for a rapid identification of the $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)$ correlation of these amino acids. As shown in Fig. 3.25, no $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)$ correlation of Ala, Val, Ser, Gly, and Phe was observed in the 2D NCA spectrum of $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils (Fig. 3.25). In contrast, the $\text{C}\alpha$ of Leu is only ^{13}C enriched by the $[1\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme. Thereby, the $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)$ correlation of Leu should be expected only in the 2D NCA spectrum of $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils (Fig. 3.25). In fact, two extra cross peaks were observed in Fig. 3.25. However, the unambiguous assignment of Leu38 of $\text{m}\alpha\text{S}$ fibrils could not be obtained so far due to a lack of sequential connections.

In total, an in-depth analysis of the labeling patterns and the characteristic correlations expected in ssNMR spectra for the 20 standard amino acids based on protein expression in a *E. coli* medium containing [2-¹³C]Glc or [1-¹³C]Glc as the sole carbon source, are demonstrated in the section 3.3 of Part I. The analysis of individual isotopomers for individual amino acids (Fig. 3.7-9 and Fig. 3.18-20) enables us to investigate to what extent the labelings are characteristic from other individual amino acids, and can thus be used for the rapid identification of special spin systems. The analysis of the ¹³C labels distribution of the 20 standard amino acids (Fig. 3.7-10 and Fig. 3.18-21), affords valuable information for the unambiguous sequential resonance assignment by homonuclear ¹³C(i)-¹³C(i±1) correlations, heteronuclear ¹⁵N(i)-¹³Cα(i) and ¹⁵N(i)-¹³C'(i-1) correlations, and will potentially provide support for collecting medium- and long-range distance correlations.

Taken together, the near-complete sequential resonance assignment for residues from Gly41 to Val95 is obtained with all the recorded spectra on [U-¹³C]Glc-, [2-¹³C]Glc- and [1-¹³C]Glc-labeled mαS fibrils. In total, 96% of backbone amide ¹⁵N and 93% of all ¹³C resonances (97% of the backbone and 89% of the sidechains) are obtained for the detected residues from Gly41 to Val95 (Table A2 in appendix A). To the best of our knowledge, our study reports the sequential resonance assignment of mαS fibrils for the first time, and with more complete information compared to hαS fibrils.

may thus be too weak to be observed. On the other hand, there are clearly no detectable sequential connections from Val95 to Lys96 in our spectra. However, we were able to find residues outside the identified fibril core in ^{13}C - ^{13}C PDS spectra, such as Ser (Ser9 or 129), Pro (Pro108, 117, 120, 128, or 138), and Asp (Asp 2, 98, 115, 119, or 135). However, we could not detect any sequential connections to assign them unambiguously. Importantly, the Ser and Pro resonances are absent in a DREAM spectrum (Fig. 3.26), which may indicate that they are relatively flexible. This is corroborated by the observation that the Ser and Pro signals have random coil chemical shifts, whereas the unassigned Asp exhibits β -strand-like chemical shifts (See Table A2 in appendix A).

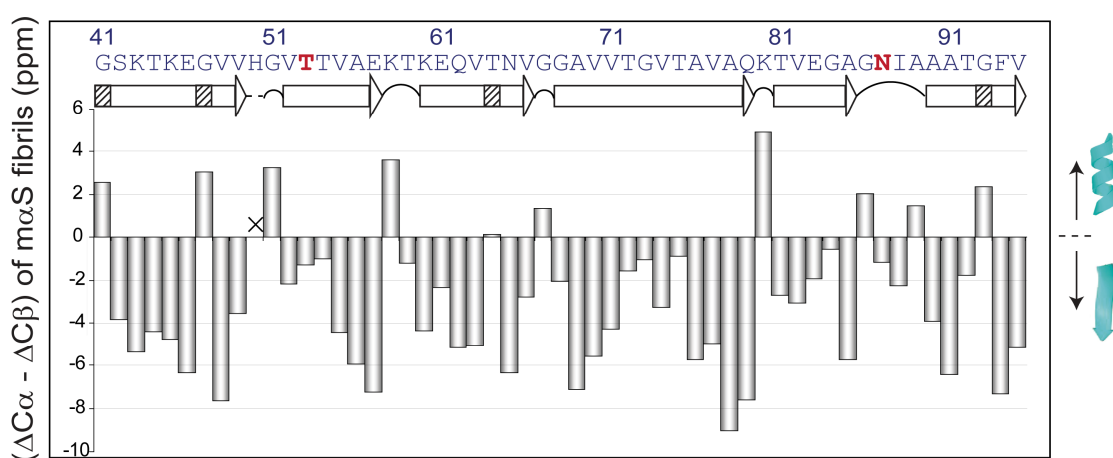


Figure 3.27. The secondary structure of the rigid core of $m\alpha S$ fibrils. Secondary chemical shifts ($\Delta C\alpha - \Delta C\beta$) are shown as a function of residue for $m\alpha S$ fibrils as studied here. Non-assigned residues are marked with a cross. β -strands are indicated by white arrows (slashed bars in the white arrow represent possible kinks), non- β -strand regions (kink, loop, or turn) are shown as a curve, and non-assigned amino acids as a dashed line. The Ala53Thr and Ser87Asn variant sites in the rigid core are colored in red.

Confident and complete sequential assignment of residues from Gly41 to Val95 leads to valuable secondary structural information. The secondary structure of the rigid core of $m\alpha S$ fibrils was analyzed based on secondary chemical shifts ($\Delta C\alpha - \Delta C\beta = [\Delta C\alpha(\text{obs}) - \Delta C\alpha(\text{rc})] - [\Delta C\beta(\text{obs}) - \Delta C\beta(\text{rc})]$) [104; 105; 106] and is illustrated in Fig. 3.27. For all the residues apart from Gly, three or more continuous negative values of ($\Delta C\alpha - \Delta C\beta$) are indicative of β -strand conformation, while positive values point to α -helical structure, and values close to zero indicate a random coil conformation.

As an independent and complementary indicator, backbone ^1H - ^1H distances were probed indirectly using the NHC scheme [51] to characterize the secondary structure of $m\alpha S$ fibrils (Fig. 3.28). Longitudinal ^1H - ^1H mixing times of 150 and 175 μs for

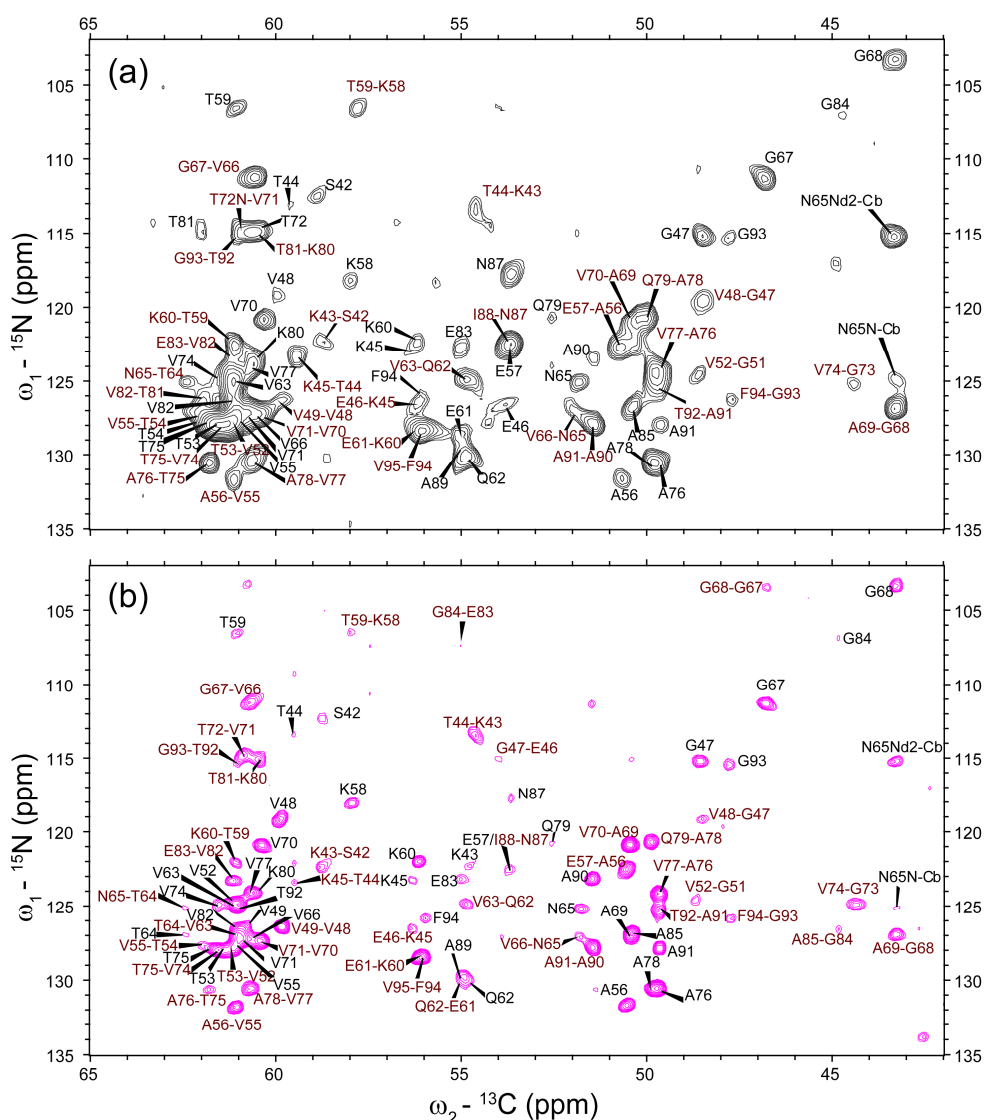


Figure 3.28. Secondary structure analysis using NHC spectra recorded on (a) [U- ^{13}C]Glc-labeled m α S fibrils and (b) [2- ^{13}C]Glc-labeled m α S fibrils. The $^{15}\text{N}(i)-^{13}\text{C}\alpha(i)$ and $^{15}\text{N}(i)-^{13}\text{C}\alpha(i-1)$ correlations are labeled in black and dark red, respectively. Both spectra were recorded at 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz. Both spectra were zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (45° shift) along both dimensions.

[U- ^{13}C]Glc- and [2- ^{13}C]Glc-labeled m α S fibrils, respectively, were used. The enhanced resolution of the NHC spectrum of [2- ^{13}C]Glc-labeled m α S fibrils compared to the spectrum of [U- ^{13}C]Glc-labeled m α S fibrils facilitated the site-specific identification of secondary structure. As illustrated in Fig. 1.2b, for residues with a β -strand conformation, inter-residue $^1\text{HN}(i)-^1\text{HC}\alpha(i-1)$ contacts are shorter than intra-residue $^1\text{HN}(i)-^1\text{HC}\alpha(i)$ contacts. Thereby, observed $^{15}\text{NH}(i+1)-^{13}\text{C}\alpha\text{H}(i)$ correlations with stronger intensities compared to $^{15}\text{NH}(i)-^{13}\text{C}\alpha\text{H}(i)$ correlations or the exclusive presence of $^{15}\text{NH}(i+1)-$

$^{13}\text{C}\alpha\text{H}(i)$ correlations, as seen e.g., for Ser42 and Lys43 are indicative of β -sheet conformation. In contrast, the clear absence of $^{15}\text{NH}(i+1)$ - $^{13}\text{C}\alpha\text{H}(i)$ correlations or the presence with weaker intensities compared to $^{15}\text{NH}(i)$ - $^{13}\text{C}\alpha\text{H}(i)$ correlations, e.g., observed for Lys58 and Thr59, is indicative of loop or turn structure. Together with the secondary chemical shifts, six β -strands were identified to be within the fibril core of αS fibrils: Gly41-Val49, Val52-Glu57, Lys60-Val66, Gly68-Gln79, Thr81-Ala85, and Ala90-Val95. Possible kinks in the β -strands are marked by slashed bars (Fig. 3.27).

3.5 Structural comparison of mouse and human α -synuclein fibrils

A direct comparison between m α S and h α S fibrils has to take into account the existence of multiple sets of chemical shifts, incomplete sequential resonance assignments, and the presence of different morphologies of h α S fibrils. E.g., 48 residues (form A, twisted fibrils) and 36 residues (form B, straight fibrils) from Leu38 to Val95 were assigned and identified to be within the fibril core of h α S for two distinct fibril morphologies by Heise et al. [44]. Recently, sequential resonance assignments for 85% of residues from Met1 to Lys97 were achieved by Gath et al. [82] and 91% of the ^{15}N and ^{13}C resonances of residues from Leu38 to Lys96 were reported by Comellas et al. [81]. Collectively, h α S fibrils seem to recruit more residues into the β -sheet fibril core than m α S fibrils. Residues from Gly41 to Val95 were sequentially assigned and identified to be within the fibril core of m α S fibrils by our current work.

The B-Form h α S fibrils studied by Heise et al. are straight, while the A-Form has a twisted morphology [44]. However, the fibril morphology was not reported in the publications by Gath et al. [82] and Comellas et al. [81] on h α S fibrils. Because the ssNMR spectra are virtually identical for h α S fibrils as reported by Gath et al. [82], and by us [91] (red spectrum in Fig. 3.2, straight fibrils, produced under the same conditions as the m α S fibrils with straight morphology reported in the current study), we assume that Gath et al. also obtained straight fibrils. However, no EM micrograph is shown in Gath et al. [82]. The fibril morphology for Comellas et al. is not clear. In the published EM micrograph (Comellas et al. [81], Fig. 8), the fibrils appear straight as well. Interestingly, we see strong differences between m α S and h α S in terms of secondary chemical shifts and extent of β -strand regions if we consider h α S prepared under the same fibrillization conditions as m α S (with assignments for this form reported by Gath et al. [82], see Fig. 3.29b). The most significant difference relates to residues Met1-Val37 that are only observed in spectra of h α S. In contrast, residues Thr44-Glu57 are only observed for m α S. Furthermore, for several residues in the core, secondary chemical shifts ($\Delta\text{C}\alpha - \Delta\text{C}\beta$) with opposite sign for m α S and h α S were observed: Gly41, Lys58, Thr81, Ala89, and Ala91. In contrast, much higher similarity was observed between m α S fibrils and h α S fibrils studied by Comellas et al. (Fig. 3.29c) [81], regarding both the length and the location of the β -strands. The only residues that adopt a significantly different secondary structure

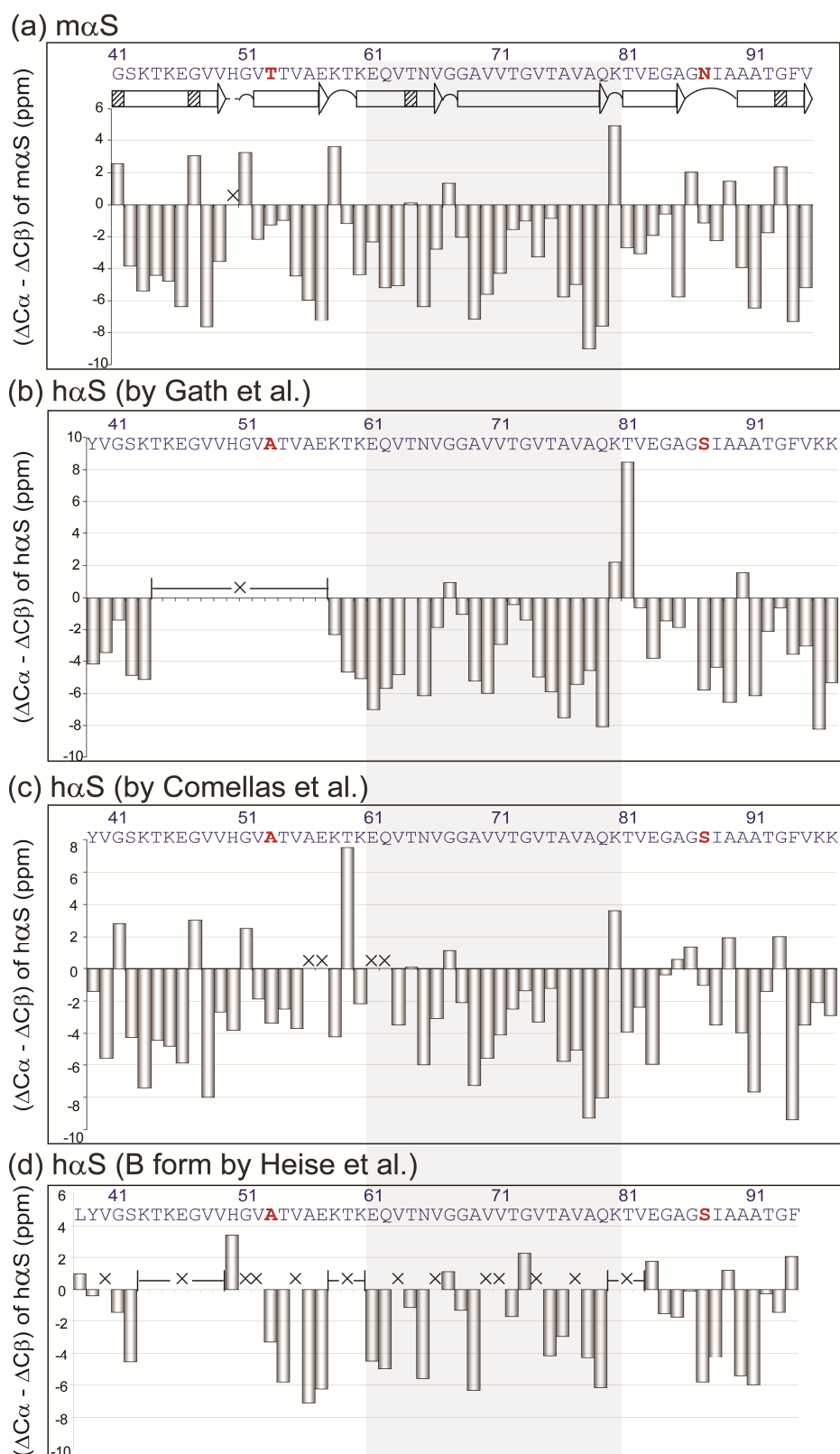


Figure 3.29. Secondary structure comparison of $m\alpha S$ fibrils and $h\alpha S$ fibrils. Secondary chemical shifts ($\Delta C\alpha - \Delta C\beta$) are shown as a function of residue for (a) $m\alpha S$ fibrils as studied here, and $h\alpha S$ fibrils as investigated by (b) Gath et al. [82] (for the sake of clarity the values for residues Met1-Leu38 are omitted), (c) Comellas et al. [81], and (d) Heise et al. [44] (only the B form of $h\alpha S$ fibrils is shown). For $h\alpha S$ fibrils, no $^{13}C\beta$ assignment was available for Val66, and Phe94 (in b), and His50, Lys60, Lys96, and Lys97 (in c). Non-assigned residues are marked with a cross. β -strands are indicated by white arrows (slashed bars in the

white arrow represent possible kinks), non- β -strand regions (kink, loop, or turn) are shown as a curve, and non-assigned amino acids as a dashed line. The Ala53Thr and Ser87Asn variant sites in the rigid core are colored in red. The conserved region of α S fibrils as revealed by the comparison of m α S and h α S fibrils studied by Gath et al. [82] and Comellas et al. [81] is highlighted in grey.

(i.e., opposite sign of the secondary chemical shift) in m α S and h α S fibrils studied by Comellas et al. [81] (Fig. 3.29c) are Lys58, Thr59, and Ala85.

The systematic structural comparison between m α S fibrils and the two different forms of h α S fibrils reveals that residues Glu61-Lys80 constitute the structurally most conserved region of α S. Interestingly, the mutation sites of Ala53Thr and Ser87Asn that are responsible for the differences in aggregation kinetics and lag phases between m α S and h α S are outside this region. The conserved core of Glu61-Lys80 might play an important role in the initiation of the aggregation of α S fibrils. Consistent with this notion, it was previously reported that residues Glu61-Ala78 [107], Gly68-Ala78 [108], and Val71-Val82 [109] within the hydrophobic region of α S could form aggregates by themselves *in vitro*, and that residues Gly68-Ala78 [108] constitute the shortest fragment sufficient for α S fibril assembly.

3.6 Supra-molecular structure of mouse α -synuclein fibrils elucidated by means of mixed 1:1 ^{13}C : ^{15}N labeled mouse α -synuclein fibrils

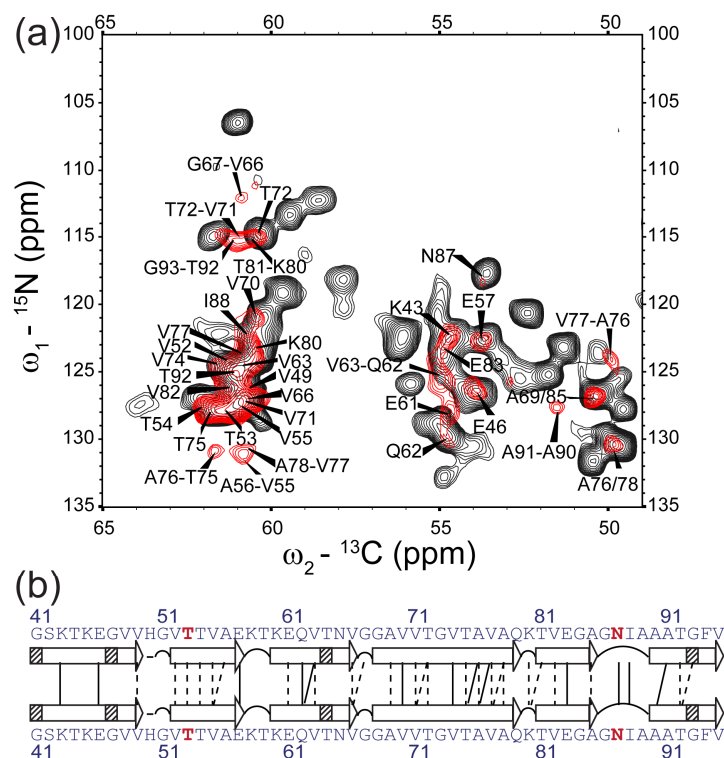


Figure 3.30. Supra-molecular arrangement of $m\alpha S$ fibrils. (a) Comparison of NCA spectrum (in black) recorded on $[U-^{13}\text{C}]$ Glc-labeled $m\alpha S$ fibrils and NHHC spectrum (in red) recorded on $[M-^{13}\text{C}/^{15}\text{N}]$ -labeled $m\alpha S$ fibrils. $^{15}\text{N}(i)-^{13}\text{C}\alpha(i)$ correlations and $^{15}\text{N}(i)-^{13}\text{C}\alpha(i-1)$ correlations between adjacent molecules were labeled as e.g., Val63 and Val63-Gln62, respectively. (b) Summary of the detected inter-molecular contacts. Solid lines represent unambiguous assignments and dashed lines represent ambiguous assignments due to spectral overlap. The red spectrum was measured on 20.0 Tesla (^1H resonance frequency: 850 MHz) and at a spinning frequency of 11 kHz, and then zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (60° shift) along both dimensions. The red spectrum was measured on 18.8 Tesla (^1H resonance frequency: 800 MHz) and at a spinning frequency of 11 kHz, and then zero-filled to 2048 (direct dimension) and 1024 (indirect dimension) points, and apodized with a squared sine bell window function (51.4° shift) along both dimensions.

Previous ssNMR studies on heterogeneously labeled samples have yielded structural insight into the supra-molecular arrangement of amyloid fibrils [29; 90; 110; 111; 112; 113; 114; 115; 116]. Here, we have recorded an NHHC [51] spectrum with a $^1\text{H}-^1\text{H}$ mixing time of 500 μs on $[M-^{13}\text{C}/^{15}\text{N}]$ -labeled $m\alpha S$ fibrils to elucidate the stacking of molecules in the fibrils. With the chosen mixing time, cross peaks between ^{13}C and ^{15}N sites can be observed that correspond to intermolecular ($^1\text{H}-^1\text{H}$) distances of about 3.5 \AA [90]. Since ^{15}N and ^{13}C spins in the $[M-^{13}\text{C}/^{15}\text{N}]$ -labeled $m\alpha S$ fibrils are not labeled

within the same molecule, all the observed cross peaks in the NHHC spectrum are exclusively due to intermolecular transfer. As shown in Fig. 3.30a, most of the cross peaks can be assigned to either inter-strand $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i)$ or inter-strand $^{15}\text{N}(i)\text{-}^{13}\text{C}\alpha(i-1)$ correlations. These contacts are summarized in Fig. 3.30b. The sensitivity of the intermolecular NHHC experiment is compromised due to a) the three involved CP transfer steps, b) the probability of one pair of neighboring molecules to be $^{15}\text{N}\text{:}^{13}\text{C}$ labeled, which is only 25%, c) the long intermolecular distances involved, and d) the distribution of magnetization during the $^1\text{H}\text{-}^1\text{H}$ transfer to protons that are not involved in the last transfer step. Still, the unambiguous intermolecular correlation pattern, e.g., for Lys43, Glu46, and Glu57 (Fig. 3.30b), as well as the overall good overlay with the NCA spectrum recorded on $[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled $\text{m}\alpha\text{S}$ fibrils (Fig. 3.30a) identifies a parallel, in-register arrangement of the β -strands. This result is in agreement with a similar recent study by our group which showed that the $\text{h}\alpha\text{S}$ fibrils studied by us are stacked parallel and in-register [91].

4 Conclusions

We have provided a complete labeling pattern for the [1-¹³C]Glc- and [2-¹³C]Glc-labeled proteins based on the major metabolic pathways. The two labeling schemes are quasi complementary, as illustrated with the absence of cross peaks detectable in both types of spectra. Substantial improvements in terms of increased spectral resolution and simplicity for [1-¹³C]Glc- and [2-¹³C]Glc-labeled m α S fibrils were observed, and strategies for simplified and complete resonance assignment using the ¹³C spin dilution approach were demonstrated with the *de novo* assignment of m α S fibrils. An almost complete resonance assignment (95% of backbone amide ¹⁵N and 92% of ¹³C nuclei) was obtained for residues from Gly41 to Val95 which form the core of m α S fibrils. Six β -strands were identified to be within the fibril core of m α S based on a secondary structure analysis.

The structural comparison between m α S fibrils and two different forms of h α S fibrils reveals that residues Glu61-Lys80 constitute the structurally most conserved region of α S. While the Glu61-Lys80 conserved core appears to play a critical role as an initiator of aggregation, a subtle modulatory effect might be exerted on the α S aggregation pathway by mutations located inside the extended fibrillogenic region formed by residues Gly41-Val95. Indeed, it was reported recently that the Ala53Thr substitution dominates the aggregation growth rates, while the combination of the Ala53Thr and Ser87Asn mutations affects the lag phase of the aggregation of m α S compared to h α S [85].

It should be worthwhile to determine structures of m α S and h α S fibrils in order to gain further insights into the largely different aggregation kinetics. Currently, we are exploiting the improved resolution and reduced dipolar truncation in the sparsely labeled m α S samples, to collect long-range distance restraints with the aim of an atomic model of m α S amyloid fibrils.

**Part II Stereospecific Assignment of Val and Leu
Prochiral Methyl Groups**

1 Introduction

The stereospecific assignment of prochiral methyl groups of Val and Leu can greatly improve the precision and accuracy in protein structure determination by solution NMR [117; 118; 119; 120]. The methyl groups of Leu and Val are useful sources of structural information, as they are often found abundantly in the structurally important hydrophobic protein core and involved in numerous inter-residue and long-range contacts [121]. Additionally, it was recently shown that the methyl carbon chemical shifts can be used for obtaining χ_1 rotamer distributions in Val and Leu residues, and that the methyl groups are a rich source of information about sidechain conformation and dynamics [122; 123]. Solid-state NMR (ssNMR) spectroscopy is an emerging powerful tool for the structure determination of noncrystalline and insoluble proteins [6; 10; 13; 29; 124], which are not amenable to X-ray crystallography or solution NMR spectroscopy. The stereospecific assignment of the prochiral methyl groups of Val and Leu by ssNMR, should play an equally important role for structural and dynamical studies as in solution NMR.

In order to address the stereospecific assignment of the prochiral methyl groups of Val and Leu, several isotope-labeling approaches have been developed [118; 119; 125; 126; 127; 128; 129]. The most widely used method to date relies on the use of a minimal culture medium containing 10% [^{13}C]Glucose and 90% [^{12}C]Glucose as the carbon source during biosynthetic protein production [118; 130; 131], and the knowledge that the biosynthesis of Val and Leu from glucose is stereoselective [118; 130; 131; 132]. As shown in Fig. 1.1, the isopropyl group is composed of a two-carbon fragment from one pyruvate unit, while the second methyl group is migrated from another pyruvate unit. This methyl migration has been shown to be stereoselective, and the migrating methyl group becomes the Pro-S methyl group in both Val and Leu (γ_2 in Val and δ_2 in Leu) [133; 134; 135; 136; 137]. Since the Pro-R methyl group (γ_1 in Val and δ_1 in Leu) and the adjacent $>\text{CH}-$ group originate from the same pyruvate molecule, they are thus ^{13}C labeled simultaneously with a probability of about 10%. In contrast, the Pro-S methyl group (γ_2 in Val and δ_2 in Leu) and the adjacent carbon atom originate from two different pyruvate molecules, and the probability of being ^{13}C labeled simultaneously is thus about 1%. The stereospecific distinction of the two isopropyl methyl groups could be easily

obtained by recording a ^1H -decoupled ^{13}C solution NMR spectrum, where the ^{13}C resonance of the Pro-R methyl group is a doublet with a splitting due to the one-bond ^{13}C - ^{13}C J coupling with the neighboring ^{13}C spin, while the ^{13}C NMR signal of the Pro-S methyl group is expected to be a singlet.

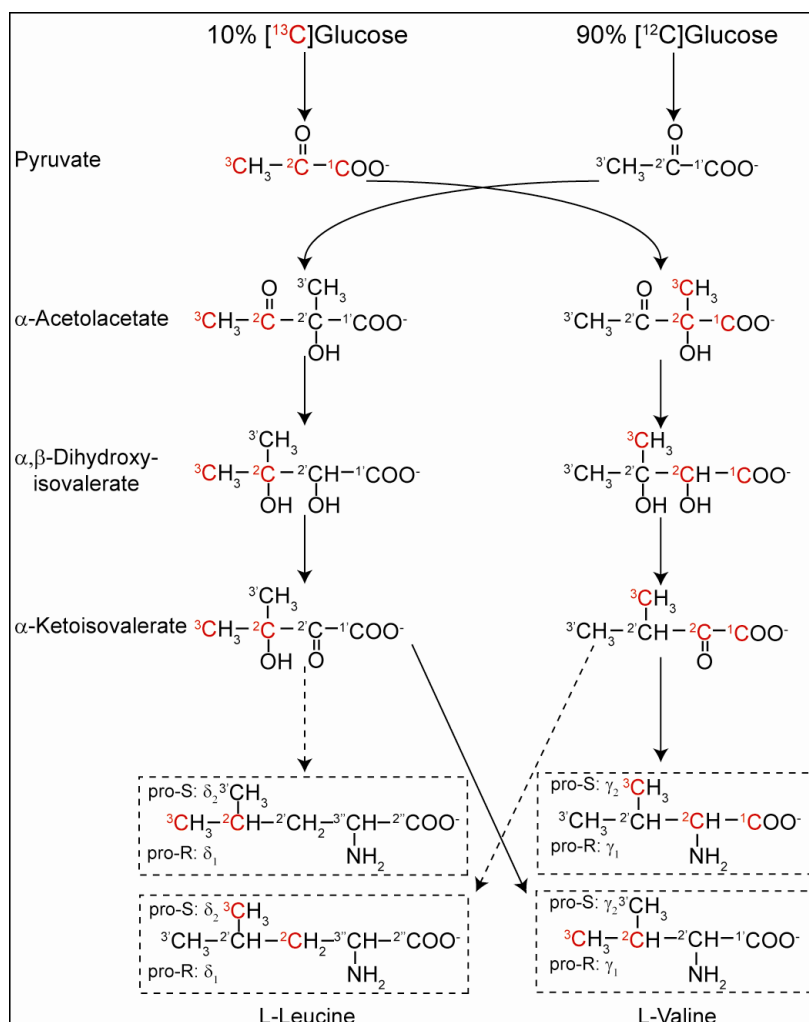


Figure 1.1. Reaction pathways for the biosynthesis of Val and Leu from a mixture of 10% [^{13}C]Glucose and 90% [^{12}C]Glucose, showing the stereochemistry and the principal labeling patterns. The carbons that are ^{13}C labeled are colored in red. For simplicity, some steps are not included, which are indicated by dashed arrows. The Figure is remade from Fig. 1 in Ref. [131].

This elegant method has been successfully applied for obtaining the stereospecific assignment of the isopropyl groups of Val and Leu by solution NMR and has also been extended to ssNMR [132]. The Pro-R methyl group and the adjacent $>\text{CH}$ - group (i.e., $\text{C}\gamma_1\text{-C}\beta$ in Val and $\text{C}\delta_1\text{-C}\gamma$ in Leu) originate from the same pyruvate molecule and are present in ^{13}C labeled pairs, while the Pro-S methyl group and the $-\text{CH}_2-$ group two bonds apart ($\text{C}\gamma_2\text{-C}\alpha\text{-C}'$ in Val and $\text{C}\delta_2\text{-C}\beta$ in Leu) originate from the same pyruvate molecule

(Fig. 1.1). Thus, the ^{13}C signals from the two prochiral methyl groups in Val and Leu can be easily distinguished by the presence or absence of a ^{13}C neighbor. In a 2D ^{13}C - ^{13}C correlation ssNMR spectrum, strong $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma 1$ / $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma 2$ in Val and $^{13}\text{C}\gamma$ - $^{13}\text{C}\delta 1$ / $^{13}\text{C}\beta$ - $^{13}\text{C}\delta 2$ in Leu are expected, while weak $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma 1$ / $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma 2$ in Val and $^{13}\text{C}\beta$ - $^{13}\text{C}\delta 1$ / $^{13}\text{C}\gamma$ - $^{13}\text{C}\delta 2$ in Leu are expected [132]. However, a major practical drawback of this method is that the fractional labeling requires an additional sample to be prepared only for the purpose of stereospecific assignments.

Here, we propose an alternative approach for the stereospecific assignment of the methyl groups of Val and Leu based on the $[2\text{-}^{13}\text{C}]\text{Glucose}$ ($[2\text{-}^{13}\text{C}]\text{Glc}$) labeling scheme. Bacterial growth in medium containing $[2\text{-}^{13}\text{C}]\text{Glc}$ results in high ^{13}C spin dilution in the protein produced with only one out of six carbons labeled [61; 89; 138]. As discussed in the Part I of this thesis, the $[2\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme leads to a significant resolution enhancement in ssNMR spectra and also to improvements in polarization transfer efficiencies due to the reduction of dipolar truncation effect [89; 91; 138]. In combination with uniformly $[^{13}\text{C}]\text{Glc}$ labeled ($[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled) and $[1\text{-}^{13}\text{C}]\text{Glc}$ -labeled samples, sequential resonance assignments can be readily obtained [89; 138]. In the Part II of this thesis, we present that the $[2\text{-}^{13}\text{C}]\text{Glc}$ labeling scheme additionally allows for obtaining stereospecific assignments of the isopropyl groups of Val and Leu in a straightforward manner using ssNMR spectroscopy. As an illustration, we present stereospecific assignments of Val and Leu in type 3 secretion system (T3SS) PrgI needles and microcrystalline ubiquitin.

2 Experiments and methods

2.1 Sample preparation

[U-¹³C]Glc- and [2-¹³C]Glc-labeled T3SS PrgI needles were prepared as described in (Loquet et al., Nature 2012) [124]. [U-¹³C]Glc- and [2-¹³C]Glc-labeled ubiquitin was prepared as described in (Seidel et al., ChembioChem 2005) [139]. [2-¹³C]Glc-labeled PrgI protomers were prepared as described in (Loquet et al., Nature 2012) [124] with the exception that the monomers were not polymerized. The buffer used for the solution NMR studies contained 20 mM MES, pH 5.5, 10% D₂O. All samples were uniformly ¹⁵N-labeled by supplementing the media with ¹⁵N-NH₄Cl as the sole nitrogen source.

2.2 Solid-state NMR experiments and data processing

Samples were packed in 4-mm MAS rotors, using protein quantities of ~10 mg for T3SS PrgI needles and ~20 mg for ubiquitin. All spectra were recorded at a spinning frequency of 11 kHz and the ¹³C chemical shifts were calibrated with DSS as an internal reference [94]. The temperature-dependent position of the water proton resonance was used to measure the temperature inside the MAS rotor [140]. High-power ¹H-¹³C decoupling (SPINAL-64 [36]) with a radio-frequency amplitude of 83 kHz was applied during evolution and detection periods.

For [U-¹³C]Glc- and [2-¹³C]Glc-labeled T3SS PrgI needles, two-dimensional ¹³C-¹³C ssNMR experiments were conducted on a 20.0 Tesla (¹H resonance frequency: 850 MHz) wide-bore spectrometer (Bruker Biospin, Germany), equipped with a 4-mm triple-resonance (¹H, ¹³C, ¹⁵N) MAS probe. The experiments were conducted at a sample temperature of 278 K (+ 5 °C). ¹³C-¹³C transfer was achieved via proton-driven spin-diffusion (PDSD) with a mixing time of 50 ms for [U-¹³C]Glc- and 400 ms for [2-¹³C]Glc-labeled PrgI, respectively. These two spectra were processed with NMRpipe [141] and analyzed using CcpNmr [142].

For [U-¹³C]Glc- and [2-¹³C]Glc-labeled ubiquitin, two-dimensional ¹³C-¹³C ssNMR experiments were conducted on a 18.8 Tesla (¹H resonance frequency: 800 MHz) standard-bore spectrometer (Bruker Biospin, Germany), equipped with a 4-mm triple-

resonance (^1H , ^{13}C , ^{15}N) MAS probe. ^{13}C - ^{13}C transfer was achieved via PDSO with mixing times of 50 ms and 100 ms for $[\text{U-}^{13}\text{C}]\text{Glc-}$ and $[\text{2-}^{13}\text{C}]\text{Glc-}$ labeled ubiquitin, respectively. The spectra were processed with Topspin (Bruker Biospin, Germany) and analyzed in SPARKY version 3.1 (T. D. Goddard & D. G. Kneller, University of California).

2.3 Solution NMR experiments and data processing

A ^1H - ^{13}C HSQC spectrum was recorded on a 21.1 Tesla (^1H resonance frequency: 900 MHz) standard-bore spectrometer (Bruker Biospin, Germany), equipped with a cryogenically-cooled triple resonance probe head, with 1024 (t2) \times 1152 (t1) complex points and maximum acquisition times of 33.6 ms (t2) and 33.9 ms (t1). 32 scans were collected per indirect increment. The ^1H carrier frequency was set at the water resonance (4.67 ppm), and the ^{13}C carrier frequency was set at 45 ppm. The spectrum was zero-filled automatically, and apodized with a squared sine bell window function along t2 and t1 axes, and then Fourier transformed. The spectrum was processed with NMRpipe [141] and analyzed using SPARKY version 3.1 (T. D. Goddard & D. G. Kneller, University of California).

3 Results and discussion

3.1 Key metabolic pathways involved in the formation of pyruvate

The ^{13}C enrichment pattern of $[2-^{13}\text{C}]\text{Glc}$ has been well established by Lundström et al. [61]. To further investigate the labeling pattern and possible scrambling, we recorded a ^1H - ^{13}C HSQC spectrum of T3SS PrgI protomers that were produced with $[2-^{13}\text{C}]\text{Glc}$ as the sole carbon source. As expected, e.g., Val $^{13}\text{C}\beta$ is present in the ^1H - ^{13}C HSQC spectrum (Fig. 3.1a), this is consistent with the labeling pattern of $[2-^{13}\text{C}]\text{Glc}$ established by Lundström et al. [61]. However, as shown in Fig. 3.1, the observation of $^{13}\text{C}\gamma_{1/2}$ of Val, $^{13}\text{C}\delta_{1/2}$ of Leu, and $^{13}\text{C}\beta$ of Ala with strong intensity indicates that metabolic scrambling indeed occurred. For example, Ala C β originates from carbon 3 of pyruvate according to its biosynthetic pathway (Fig. A1a). The presence of Ala $^{13}\text{C}\beta$ in the ^1H - ^{13}C HSQC spectrum of T3SS PrgI protomers thus indicates that, e.g., $[3-^{13}\text{C}]\text{pyruvate}$ and/or $[1,3-^{13}\text{C}]\text{pyruvate}$, are present.

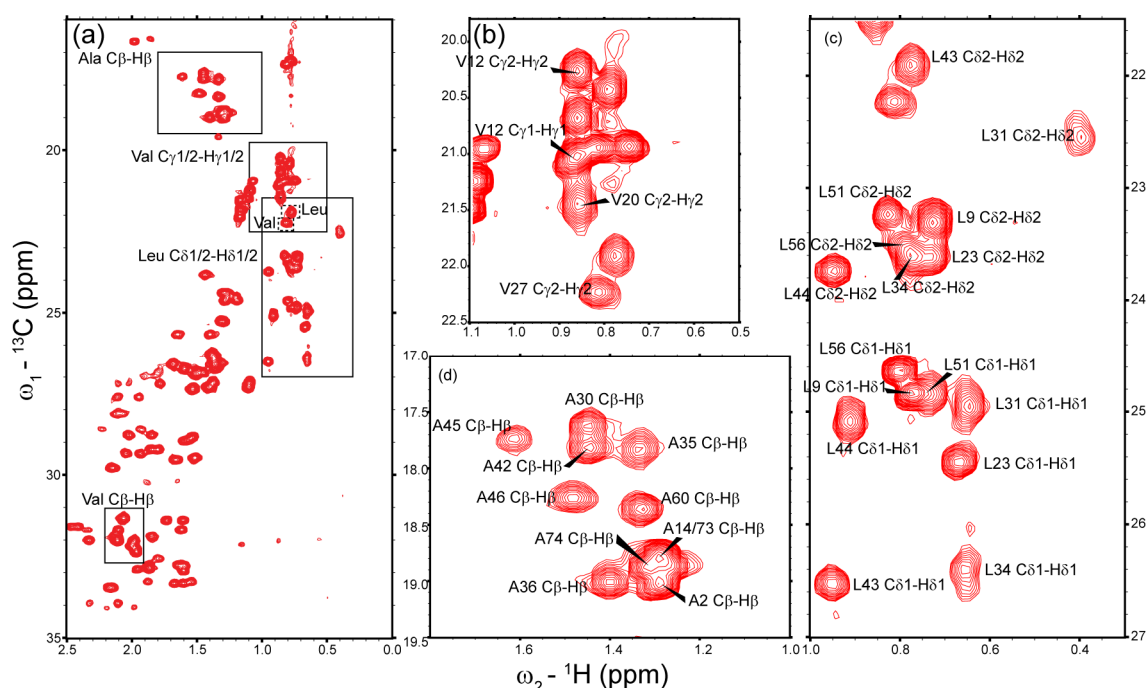


Figure 3.1. Selected region of a ^1H - ^{13}C HSQC spectrum recorded on monomeric $[2-^{13}\text{C}]\text{Glc}$ -labeled PrgI (a) and excerpts containing the correlations of (b) Val $^{13}\text{C}\gamma_{1/2}$ - $^1\text{H}\gamma_{1/2}$, (c) Leu $^{13}\text{C}\delta_{1/2}$ - $^1\text{H}\delta_{1/2}$, and (d) Ala $^{13}\text{C}\beta$ - $^1\text{H}\beta$. Partial assignments were obtained based on the assignment of the Val65Ala/Val67Ala double mutant T3SS PrgI protomer [143].

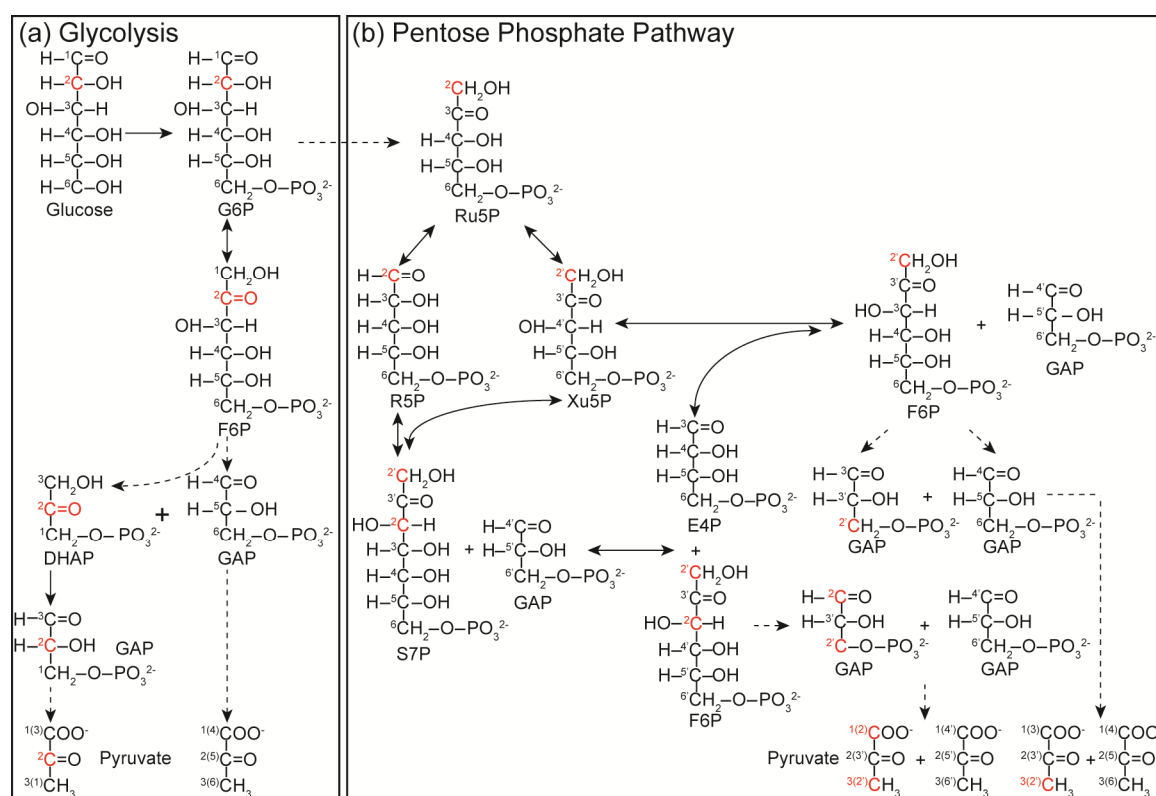


Figure 3.2. Key metabolic pathways which are involved in the formation of pyruvate when *E. coli* is grown on [2-¹³C]Glc as the sole carbon source, (a) glycolysis, and (b) Pentose phosphate pathway (PPP). The *superscript numbers* do not correspond to the nomenclature number of a given molecule but indicate the original carbon position in the six-carbon [2-¹³C]Glc chain. For example, carbon atoms labeled with ‘2’ are derived uniquely from carbon 2 of [2-¹³C]Glc, whereas carbons labeled with “2’” (primed) are derived from the carbon 2’ of a second [2-¹³C]Glc molecule. The carbons that are ¹³C labeled are colored in red. For the sake of clarity, sugars are shown in their linear forms. For simplicity, some steps are not included, which are indicated by dashed arrows. Double-headed arrows indicate reversible reactions. Abbreviations: Glucose-6-phosphate, G6P; Fructose-6-phosphate, F6P; Dihydroxyacetone phosphate, DHAP; Glyceraldehyde-3-phosphate, GAP; Ribulose-5-phosphate, Ru5P; Ribose-5-phosphate, R5P; Xylulose-5-phosphate, Xu5P; Sedoheptulose-7-phosphate, S7P; Erythrose-4-phosphate, E4P.

Starting from glucose, glycolysis leads to pyruvate as its final product (Fig. 3.2a) [100]. As a consequence, about 50% [2-¹³C]pyruvate is expected when starting from [2-¹³C]Glc. As an alternative pathway to glycolysis, the pentose phosphate pathway (PPP) [7; 100] yields three- and six carbon glycolytic intermediates, e.g., Fructose-6-phosphate (F6P) and Glyceraldehyde-3-phosphate (GAP) from [2-¹³C]Glucose-6-phosphate (G6P) by a complicated series of carbon-transfer reactions (Fig. 3.2b) depending on the metabolic needs of the cell. The detailed flow of isotope labels in PPP starting from [2-¹³C]G6P is shown in Fig2.3b. Generally, the PPP is conventionally divided into oxidative (i.e., reactions from G6P to Ru5P, and reactions 1-2 as shown in Fig. 3.2b) and non-oxidative phases (reactions 3-5 shown in Fig. 3.2b) [7]. In the oxidative phase, G6P is converted to the 5 carbon sugar R5P via a series of reactions by the loss of carbon 1 in

G6P. R5P could be isomerized to Xu5P. In the non-oxidative phase, the five carbon sugars, i.e., R5P and Xu5P, are converted to F6P and E4P or F6P and GAP by the enzymes of transketolase and transaldolase (i.e., via reactions of 3-4 or 5 in Fig 2.3b).

As a consequence, the PPP [100] yields $[1-^{13}\text{C}]/[1,3-^{13}\text{C}]$ F6P and GAP starting from $[2-^{13}\text{C}]$ G6P (Fig. 3.2b). The resulting F6P and GAP may reenter glycolysis and then yield pyruvate [100; 144], with $[1-^{13}\text{C}]$ F6P and $[1,3-^{13}\text{C}]$ F6P leading to the formation of $[3-^{13}\text{C}]$ pyruvate and $[1,3-^{13}\text{C}]$ pyruvate, respectively. Additionally, the $[2-^{13}\text{C}]$ F6P obtained during glycolysis may also be converted to $[2-^{13}\text{C}]$ R5P and $[2-^{13}\text{C}]$ Xu5P (see reactions 1-5 in Fig. 3.2), which can then be converted to $[2-^{13}\text{C}]$ F6P or $[2,3-^{13}\text{C}]$ F6P (via reactions 3-4 or 5; see Fig. 3.2) due to the rapid equilibration of the transketolase and pentose-5-phosphate isomerization reactions. Subsequently, $[2-^{13}\text{C}]$ pyruvate or $[1,2-^{13}\text{C}]$ pyruvate can be formed. As a summary, all the types of pyruvate scrambled in PPP from the combination of $[2-^{13}\text{C}]/[1-^{13}\text{C}]/[^{12}\text{C}]$ R5P and $[2-^{13}\text{C}]/[1-^{13}\text{C}]/[^{12}\text{C}]$ Xu5P (via reactions 3-4 in Fig. 3.2b), or the combination of $[2-^{13}\text{C}]/[1-^{13}\text{C}]/[^{12}\text{C}]$ Xu5P and $[1-^{13}\text{C}]/[^{12}\text{C}]$ E4P (via reaction 5 in Fig. 3.2b), are shown in appendix B Fig. B1. In total, 6 isotopomers of pyruvate are formed, $[2-^{13}\text{C}]$ pyruvate, $[3-^{13}\text{C}]$ pyruvate, $[1,3-^{13}\text{C}]$ pyruvate, $[1,2-^{13}\text{C}]$ pyruvate, $[1-^{13}\text{C}]$ pyruvate, and $[1,2,3-^{12}\text{C}]$ pyruvate.

As illustrated above, the PPP, especially the recycling action of transketolase and transaldolase results in considerable scrambling of isotope labeling patterns of pyruvate [7]. Indeed, it is believed that the PPP is the most important alternative pathway that can ultimately yields pyruvate in addition to glycolysis [145; 146]. It was also reported that, about 75% of the pyruvate molecules are synthesized via glycolysis, and about 25% of the pyruvate molecules are synthesized via the PPP [147].

3.2 Stereoselective biosynthesis of Val and Leu

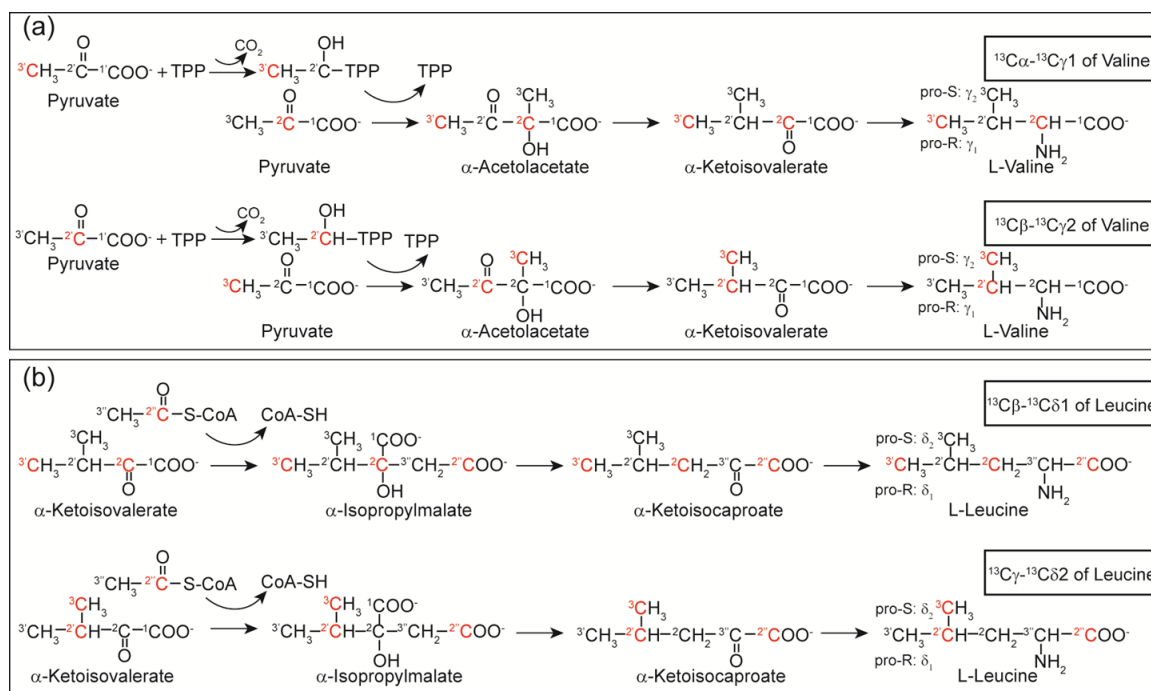


Figure 3.3. The stereoselective biosynthesis of Val and Leu. As an example, the formation of the isotope pairs (a) $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1/^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ of Val, and (b) $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_2/^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_1$ of Leu originating from $[2-^{13}\text{C}]$ pyruvate and $[3-^{13}\text{C}]$ pyruvate is shown. The carbons that are ^{13}C labeled are colored in red. Abbreviations: TPP, thiamine pyrophosphate; CoA: Coenzyme A.

Moreover, the biosynthesis of Val and Leu from pyruvate is stereoselective [118; 130; 131; 132]. As a consequence, with the presence of the above mentioned isotopomers of pyruvate, e.g., $[2-^{13}\text{C}]$ pyruvate and $[3-^{13}\text{C}]$ pyruvate, only the Pro-S methyl carbon is ^{13}C enriched simultaneously with the directly bonded carbon, while the Pro-R carbon is ^{13}C enriched simultaneously with the carbon two bonds apart ($\text{C}\alpha$ in Val and $\text{C}\beta$ in Leu). As shown in Fig. 3.3, the labeled spin pairs $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ in Val (Fig. 3.3a), and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_2$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_1$ in Leu (Fig. 3.3b) are obtained. Thus, by means of $^{13}\text{C}\text{-}^{13}\text{C}$ correlation ssNMR spectroscopy, the two prochiral methyl groups of Val and Leu can be readily distinguished and the stereospecific assignment obtained. The resulting $^{13}\text{C}\text{-}^{13}\text{C}$ correlation spectrum will exhibit strong intensity correlations for $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ of Val, and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_2$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_1$ of Leu, respectively. In contrast, the correlations $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_1$ of Val, and $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_2$ and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_1$ of Leu are not expected or with much weaker intensity. In case of a crowded methyl region in $^{13}\text{C}\text{-}^{13}\text{C}$ correlation ssNMR spectra, this method thus bears another advantage, i.e., the

simplification by the decreased number of detected correlations, in addition to the spectroscopic differentiation of the prochiral methyl groups.

3.3 Stereospecific assignment of Val and Leu in T3SS PrgI needles and ubiquitin

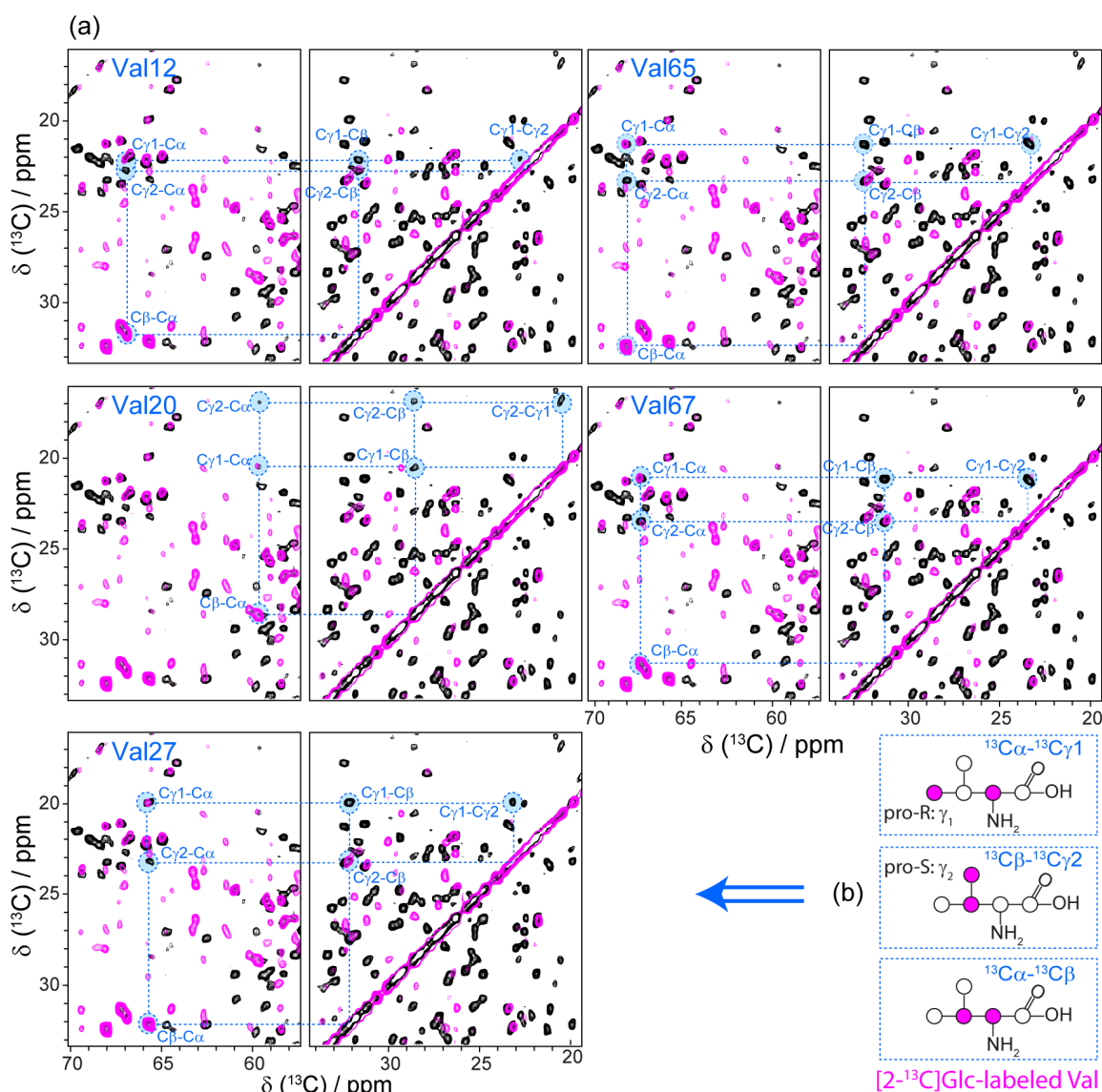


Figure 3.4. (a) 2D ^{13}C - ^{13}C PDSF spectra of $[U-^{13}\text{C}]$ Glc- (in black) and $[2-^{13}\text{C}]$ Glc-labeled (in magenta) T3SS PrgI needles. The spectrum of the $[U-^{13}\text{C}]$ Glc-labeled sample was recorded with a mixing time of 50 ms, using maximum acquisition times of 20 ms (direct dimension) and 15 ms (indirect dimension). The total experimental time was 3 days. The spectrum of the $[2-^{13}\text{C}]$ Glc-labeled sample was recorded with a mixing time of 400 ms, using maximum acquisition times of 15 ms (direct dimension) and 10 ms (indirect dimension). The total experimental time was 1.5 days. Spin systems for Val12, Val20, Val27, Val65, and Val67 are highlighted in blue. The stereospecific assignment of $^{13}\text{C}\gamma_1/2$ in Val (a) was obtained based on the correlations $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma_1$ and $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma_2$ as illustrated in (b). 1D slices for the residues Val65 and Val67 are shown in appendix B Fig. B2.

The stereospecific ^{13}C labeling pattern of Val and Leu resulting from $[2-^{13}\text{C}]$ Glc is successfully demonstrated with applications to T3SS PrgI needles (Fig. 3.4a), where

stereospecific assignments of Val and Leu are obtained unambiguously. As shown in Fig. 3.4a, generally only one methyl carbon ($^{13}\text{C}\gamma_1$ or $^{13}\text{C}\gamma_2$) correlates to $^{13}\text{C}\alpha$ or $^{13}\text{C}\beta$ of Val in the 2D ^{13}C - ^{13}C PDS spectrum of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled T3SS PrgI needles (in magenta), while four correlations ($^{13}\text{C}\gamma_1/2\text{-}^{13}\text{C}\beta$ and $^{13}\text{C}\gamma_1/2\text{-}^{13}\text{C}\alpha$) are present in the 2D ^{13}C - ^{13}C PDS spectrum of $[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled T3SS PrgI needles (in black). For Val12, 27, and 65, as expected, the correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ were observed for $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled T3SS PrgI needles (magenta spectrum in Fig. 3.4a). For Val20, only the correlation of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ was observed, but with weak intensity, while the correlation of $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ was absent. This observation is consistent with the observation of weak-intensity correlations in the 2D ^{13}C - ^{13}C PDS spectrum of $[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled T3SS PrgI needles (in black) and can be attributed to structural plasticity [124]. For Val67, in addition to the $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ correlations, a $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$ correlation was also observed, but with much weaker intensity compared to the expected correlations. Peaks that are observed in the $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled spectrum in addition to the ones observed in the $[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled sample correspond to sequential, medium- and long-range correlations, as the ^{13}C - ^{13}C PDS mixing time used was longer (400 ms).

Furthermore, the utility of this methodology is successfully demonstrated and confirmed by ubiquitin. As shown in Fig. 3.5a, only one methyl carbon ($^{13}\text{C}\gamma_1$ or $^{13}\text{C}\gamma_2$) correlates to $^{13}\text{C}\alpha$ or $^{13}\text{C}\beta$ of Val in the 2D ^{13}C - ^{13}C PDS spectrum of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled Ubiquitin (in magenta), while four correlations ($^{13}\text{C}\gamma_1/2\text{-}^{13}\text{C}\beta$ and $^{13}\text{C}\gamma_1/2\text{-}^{13}\text{C}\alpha$) are present in the 2D ^{13}C - ^{13}C PDS spectrum of $[\text{U-}^{13}\text{C}]\text{Glc}$ -labeled ubiquitin (in black). Taking Val26 as an example, as expected, only the correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ were observed in the 2D ^{13}C - ^{13}C PDS spectrum of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled ubiquitin (in magenta). However, we need to note that for Val17, in addition to the correlations of $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$, $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_1$ was also observed, but with much weaker intensity compared to its counterpart (magenta spectrum in Fig. 3.5a). Similarly for Leu, the stereoselective ^{13}C enrichment pattern, i.e., $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_1$ and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_2$, is illustrated using $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled ubiquitin (magenta spectrum in Fig. 3.5b). Taking Leu56 as an example, the correlations $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_1$ and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_2$ were observed without the presence of the correlations $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_2$ and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_1$. However, for Leu50, only the correlation

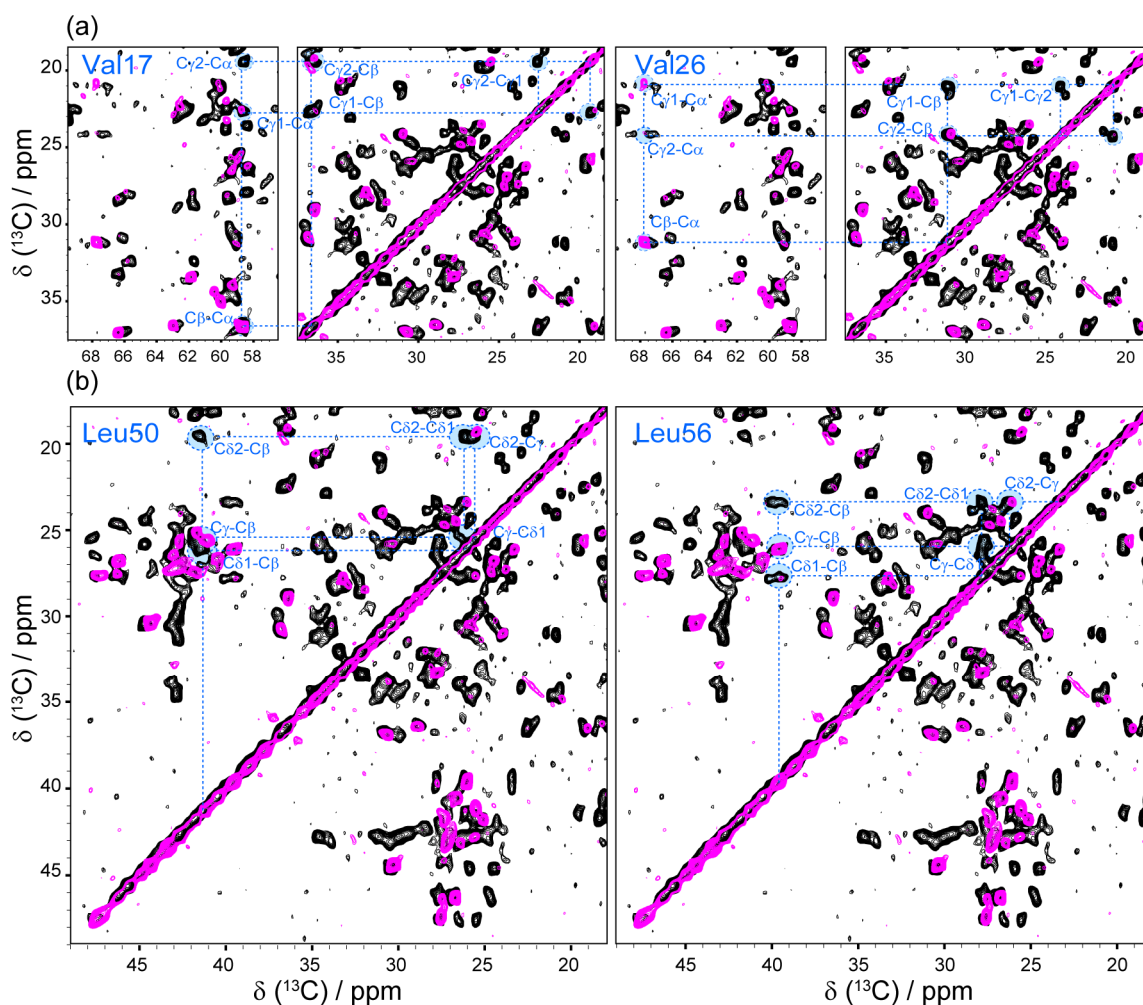


Figure 3.5. The stereospecific assignment of (a) Val and (b) Leu in ubiquitin. 2D PDS D spectra of [U- ^{13}C]Glc-labeled ubiquitin with a mixing time of 50 ms (in black) and [2- ^{13}C]Glc-labeled ubiquitin with a mixing time of 100 ms (in magenta). Spin systems for (a) V17 and V26, and (b) L50 and L56 are highlighted in blue.

of $^{13}\text{C}_\gamma$ - $^{13}\text{C}_{\delta 2}$ was clearly present, while the correlation of $^{13}\text{C}_\beta$ - $^{13}\text{C}_{\delta 1}$ was very weak and only observable on one side of the diagonal. This might be due to the longer distance involved in this correlation (two bonds) compared to the $^{13}\text{C}_\gamma$ - $^{13}\text{C}_{\delta 2}$ (one bond) correlation. Additionally, strong $^{13}\text{C}_\beta$ - $^{13}\text{C}_\gamma$ cross peaks are observed for Leu. These peaks are expected from the labeling pattern and are very intense compared to the $^{13}\text{C}_\gamma$ - $^{13}\text{C}_{\delta 2}$ and $^{13}\text{C}_\beta$ - $^{13}\text{C}_{\delta 1}$ peaks that are only observed due to scrambling.

The results on [2- ^{13}C]Glc-labeled T3SS PrgI needles and ubiquitin, show that in general the stereoselective ^{13}C enrichment pattern of Val and Leu follows the expectations and that the stereospecific assignment of the prochiral methyl carbons can be readily obtained. In principle, one can directly identify Val and Leu methyl groups by

their expected correlations in 2D ^{13}C - ^{13}C PDS spectra of $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeled samples, i.e., $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_1$ and $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ for Val, and $^{13}\text{C}\beta\text{-}^{13}\text{C}\delta_1$ and $^{13}\text{C}\gamma\text{-}^{13}\text{C}\delta_2$ for Leu. However, in case of absence of expected peaks, e.g., as seen in this study for Val20 $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_2$ in T3SS PrgI needles (Fig. 3.4a), one also needs $[U\text{-}^{13}\text{C}]\text{Glc}$ spectra to assist the assignment of $^{13}\text{C}\gamma_1/2$.

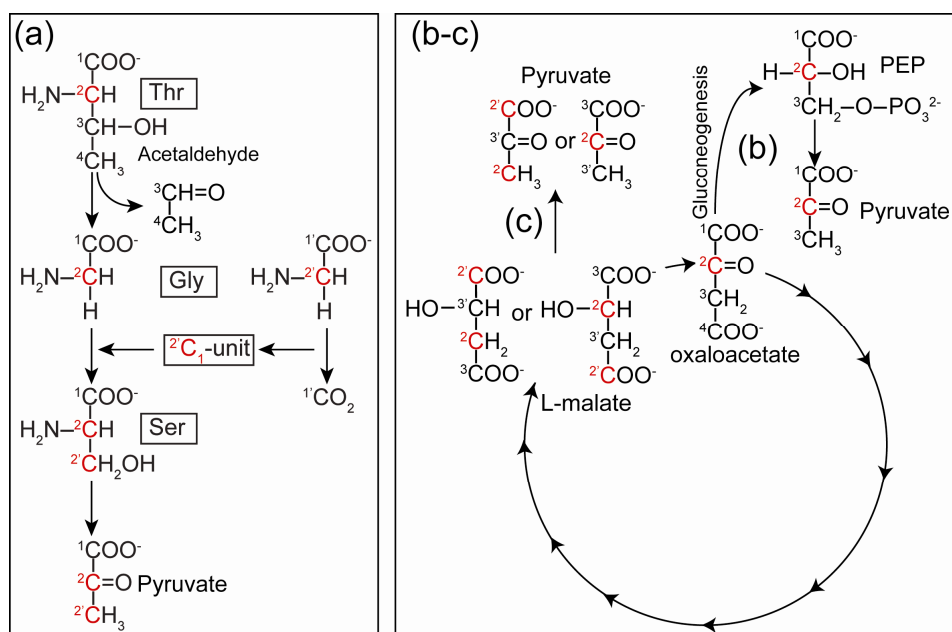


Figure 3.6. Different metabolic pathways might be involved in the scrambling of pyruvate. As an example for individual case, (a) the isotopomer of $[2,3\text{-}^{13}\text{C}]\text{pyruvate}$ derived from amino acid degradation and the C_1 metabolism, and (b) the isotopomer $[2\text{-}^{13}\text{C}]\text{pyruvate}$ derived via the breakdown of oxaloacetate, as well as (c) the isotopomers $[1,3\text{-}^{13}\text{C}]\text{pyruvate}$ and $[2\text{-}^{13}\text{C}]\text{pyruvate}$ converted from malate, are illustrated here. The carbons that are ^{13}C labeled are colored in red.

In addition to the scrambling due to the PPP detailed above, it is noteworthy that a) amino acid degradation together with the C_1 metabolism [100] (Fig. 3.6a), b) the breakdown of oxaloacetate via gluconeogenesis [100] (Fig. 3.6b) as well as c) the conversion of malate [100] (Fig. 3.6b) might contribute to scrambling resulting in a) $[3\text{-}^{13}\text{C}]\text{pyruvate}$, $[1\text{-}^{13}\text{C}]\text{pyruvate}$, $[1,3\text{-}^{13}\text{C}]\text{pyruvate}$, and $[2,3\text{-}^{13}\text{C}]\text{pyruvate}$, b) $[1,3\text{-}^{13}\text{C}]\text{pyruvate}$ and $[1\text{-}^{13}\text{C}]\text{pyruvate}$, and c) $[1,3\text{-}^{13}\text{C}]\text{pyruvate}$ and $[1\text{-}^{13}\text{C}]\text{pyruvate}$ ($[2\text{-}^{13}\text{C}]\text{pyruvate}$ is not considered here, as it is always present in the $[2\text{-}^{13}\text{C}]\text{Glc}$ -labeling scheme). As mentioned above, the spin pairs of $^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma_1$ and $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$ of Val originate from carbon 2 and 3 of a single pyruvate (Fig. 3.3). Therefore, the presence of $[2,3\text{-}^{13}\text{C}]\text{pyruvate}$ could explain the presence of the $^{13}\text{C}\alpha\text{-}^{13}\text{C}\gamma_2$ correlation of Val67 in T3SS PrgI needles (Fig. 3.4a). However, consistent with the observed weak intensity, it

was reported that amino acid degradation is quite generally of minor importance when *E. coli* cells have access to glucose [103]. This implies that the enrichment level of [2,3-¹³C]pyruvate is so low that it doesn't disturb the stereospecific assignment of Val and Leu. This is also corroborated by the fact, that no Ala ¹³C α -¹³C β peaks are observed in ssNMR spectra of [2-¹³C]Glc-labeled proteins [89; 91; 138], which would be the case if [2,3-¹³C]pyruvate was present.

4 Conclusions

We have presented an alternative stereospecific labeling pattern of Val and Leu formed by the co-presence of [2-¹³C]pyruvate and [3-¹³C]/[1,3-¹³C]pyruvate starting from [2-¹³C]Glc, enabling us to obtain stereospecific assignments of the isopropyl groups of Val and Leu in a simple way. For this purpose, only a unique [2-¹³C]Glc-labeled sample is required, the isotope expense is thus reduced compared to fractional labeling approaches [118; 130; 131; 132]. Given the resolution enhancement and spectral simplification that we have recently exploited for the *de novo* sequential assignment of the T3SS PrgI needles [89] and of mouse α -synuclein fibrils [138], the detection of long-range distance restraints [124], and the stereospecific assignment of the isopropyl groups of Val and Leu presented here, the [2-¹³C]Glc labeling scheme has proven to be a remarkably versatile scheme, and should therefore become an attractive tool for structural investigations of proteins by ssNMR.

Part III Summary

In Part I of this thesis, we have structurally characterized α S fibrils using ssNMR. The assignment strategy employed a set of high-resolution 2D and 3D ssNMR spectra recorded on [U- ^{13}C]Glc-, [1- ^{13}C]Glc-, and [2- ^{13}C]Glc-labeled α S fibrils. Proteins expressed in a *E. coli* medium containing [1- ^{13}C]Glc or [2- ^{13}C]Glc as the sole carbon source are very sparsely labeled, with only one out of six carbons being ^{13}C labeled. A detailed analysis of the labeling patterns and the characteristic correlations expected in ssNMR spectra using the [1- ^{13}C]Glc or [2- ^{13}C]Glc labeling scheme is demonstrated. The analysis affords valuable information for a rapid identification of spin systems. The resulting 2D ssNMR spectra of [1- ^{13}C]Glc- and [2- ^{13}C]Glc-labeled α S fibrils exhibit high resolution and contain a reduced number of cross peaks. This permits a near-complete backbone and sidechain resonance assignment of α S fibrils in a straightforward way. 96% of backbone amide ^{15}N and 93% of all ^{13}C nuclei were obtained for residues from Gly41 to Val95 which form the core of α S fibrils. Moreover, six β -strands are identified to be within the fibril core of α S based on a secondary chemical shift and NHHC analysis. Intermolecular ^{13}C - ^{15}N restraints obtained from [M- $^{13}\text{C}/^{15}\text{N}$]-labeled α S fibrils reveal a parallel, in-register supra-molecular β -sheet arrangement. The results were compared to recent structural studies on $\text{h}\alpha$ S fibrils and indicate the presence of a structurally conserved motif comprising residues Glu61-Lys80.

In Part II of this thesis, we have presented an alternative strategy for ssNMR stereospecific resonance assignment of the prochiral methyl groups of Val and Leu based on [2- ^{13}C]Glc as the sole carbon source during protein expression. A new stereospecific labeling pattern of Val and Leu is formed by the co-presence of [2- ^{13}C]pyruvate and [3- ^{13}C]/[1,3- ^{13}C]pyruvate starting from [2- ^{13}C]Glc, i.e., $^{13}\text{C}\beta$ - $^{13}\text{C}\gamma$ 2 and $^{13}\text{C}\alpha$ - $^{13}\text{C}\gamma$ 1 in Val, and $^{13}\text{C}\gamma$ - $^{13}\text{C}\delta$ 2 and $^{13}\text{C}\beta$ - $^{13}\text{C}\delta$ 1 in Leu. The stereospecific ^{13}C labeling pattern of Val and Leu has been successfully demonstrated with applications to T3SS PrgI needles and ubiquitin by means of ^{13}C - ^{13}C correlation ssNMR spectroscopy, where stereospecific assignments of Val and Leu are obtained unambiguously. Moreover, this method bears another advantage, i.e., spectral simplification in the often congested methyl region.

Part IV Appendix

Appendix A

Structural Comparison of Mouse and Human α -Synuclein Amyloid Fibrils by Solid-State NMR and Sparse ^{13}C Labeling Schemes

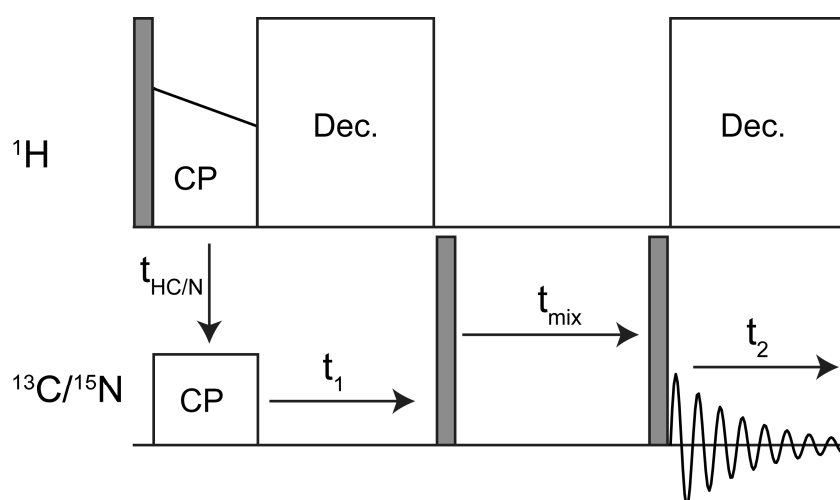


Figure A1. Pulse sequence for a 2D ^{13}C - ^{13}C or ^{15}N - ^{15}N PDS correlation spectrum. t_{mix} : spin diffusion mixing time.

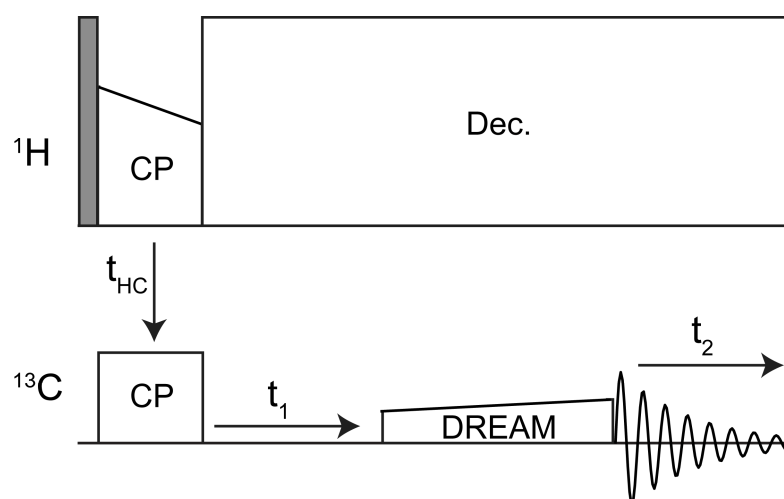


Figure A2. Pulse sequence for a DREAM [97] spectrum.

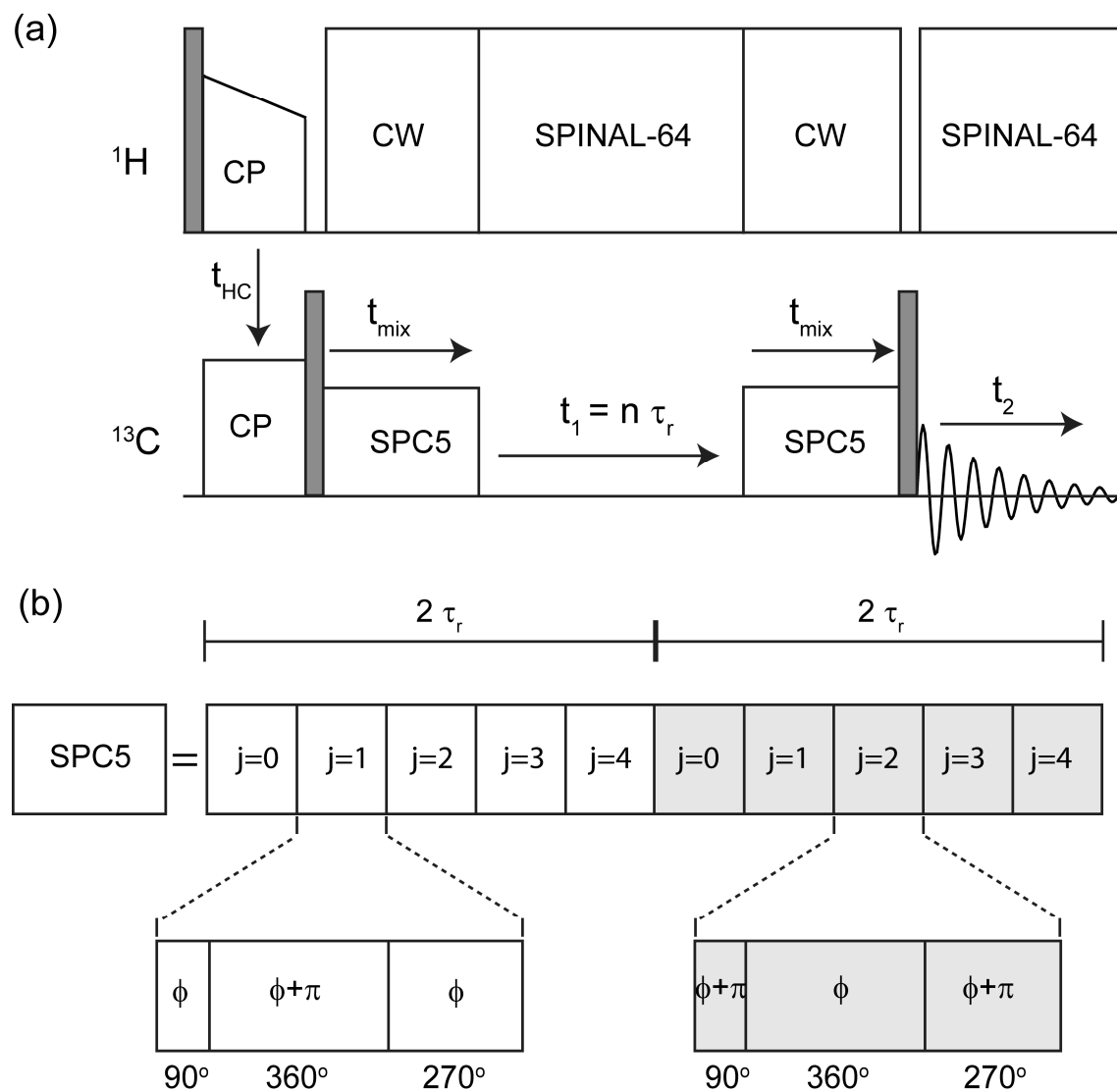


Figure A3. (a) Schematic pulse sequence of a 2D ^{13}C - ^{13}C (2Q,1Q) correlation spectrum employing SPC5 2Q excitation scheme [98]. (b) Detailed illustration of a SPC5-block, which consists of five composite pulses with a total rotation of each $(90^\circ\text{-}360^\circ\text{-}270^\circ) 0^\circ$. Individual composite pulses are repeated with shifts in their phases $\phi = j \times (2\pi/5)$. Every second SPC5 element is repeated with an additional π phase shift, indicated by shaded boxes. One SPC5 element is span over 2 rotor periods (τ_r).

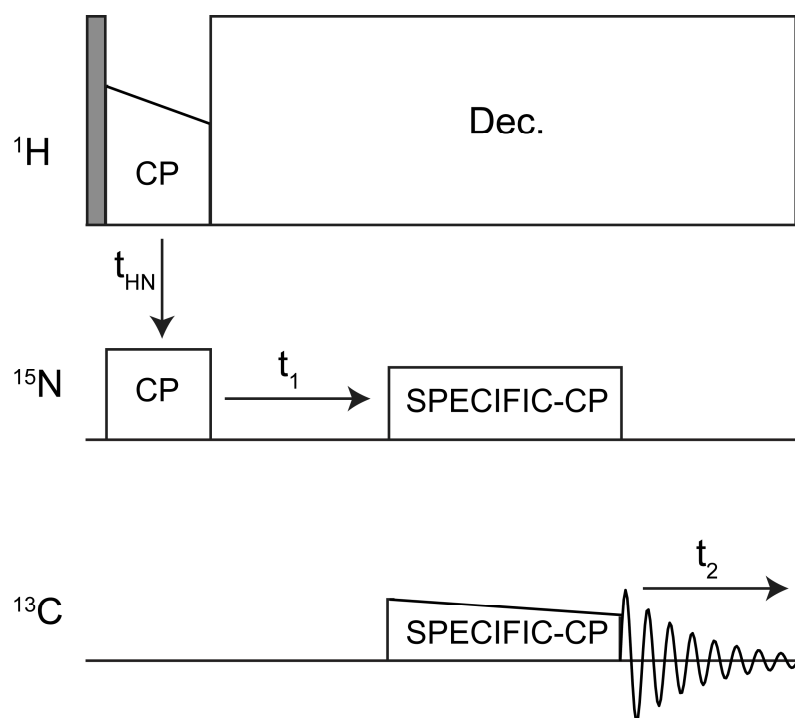


Figure A4. Pulse sequence for a 2D ^{15}N - ^{13}C correlation spectrum via SPECIFIC-CP [49; 50] to selectively transfer magnetization from backbone ^{15}N to $^{13}\text{C}\alpha$ or $^{13}\text{C}'$ nuclei, i.e., NCO and NCA.

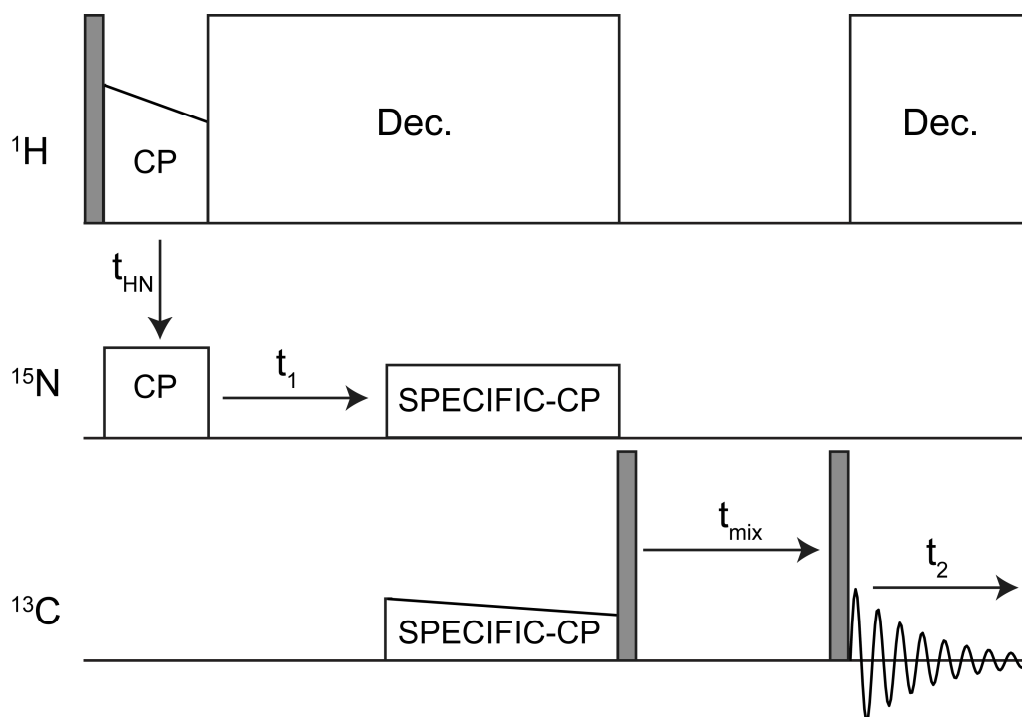


Figure A5. Pulse sequence for 2D NCACX and NCOCX experiments, the ^{13}C - ^{13}C correlation is established via PSD mixing (t_{mix}).

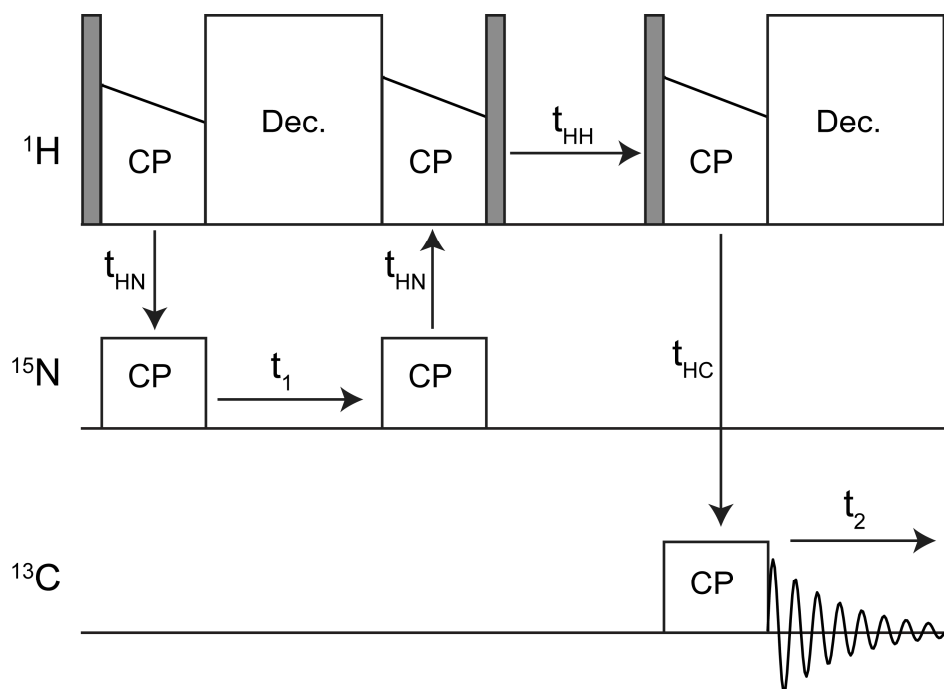


Figure A6. Experimental NHHC correlation scheme with a longitudinal mixing. Shaded rectangles in dark grey correspond to 90° nutations. Dec. indicates r.f. decoupling of protons from ^{13}C or ^{15}N spins. ^1H - ^1H contacts are encoded during a two-dimensional ^{15}N - ^{13}C correlation experiment.

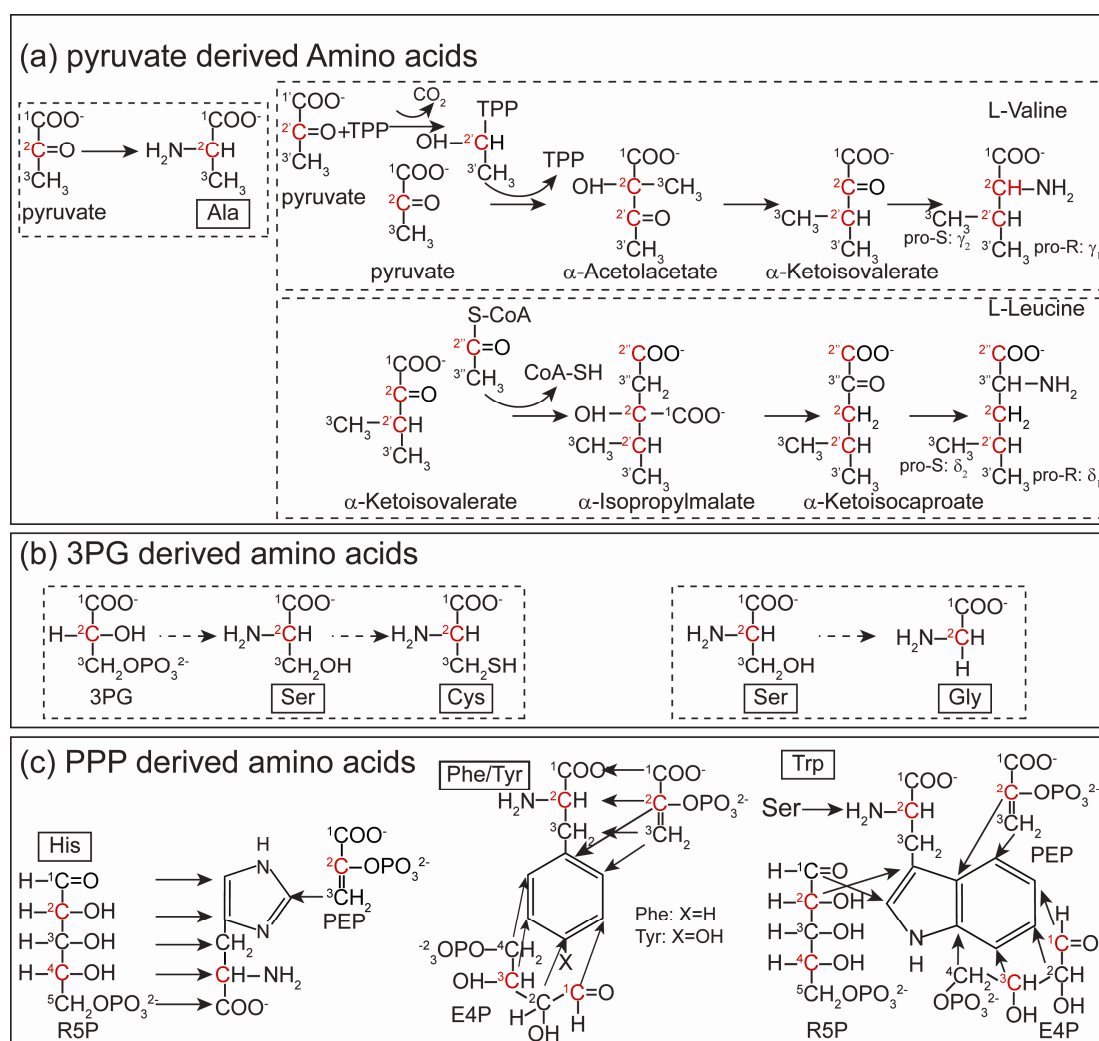


Figure A7. Biosynthetic pathways of amino acids derived from a) pyruvate, b) 3PG, and c) PPP intermediates for proteins expressed in an *E.coli* medium containing $[2-^{13}\text{C}]\text{Glc}$ as the sole carbon source. The carbons that are ^{13}C labeled are colored in red. Abbreviations: thiamine pyrophosphate, TPP; Coenzyme A, CoA.

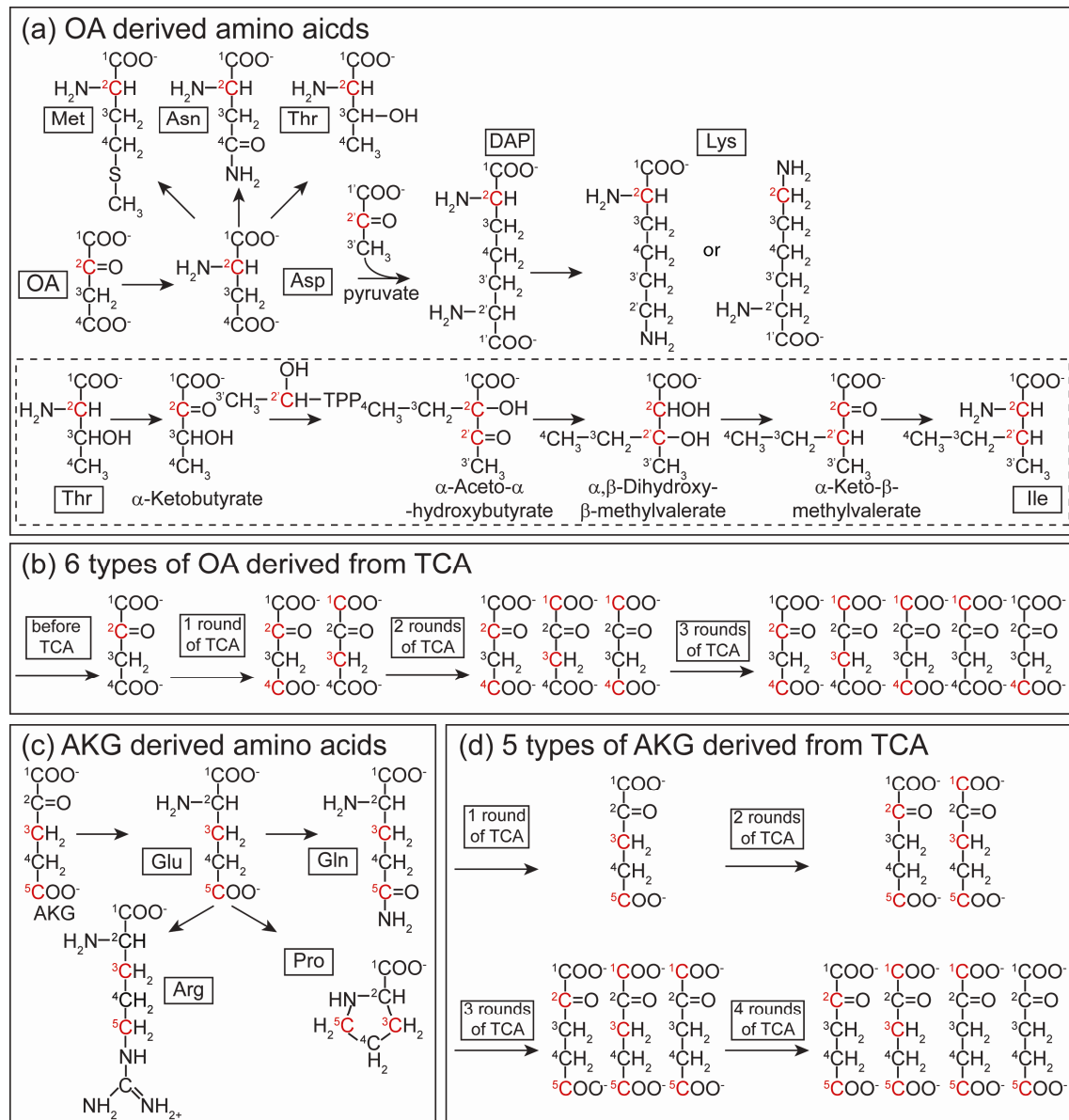
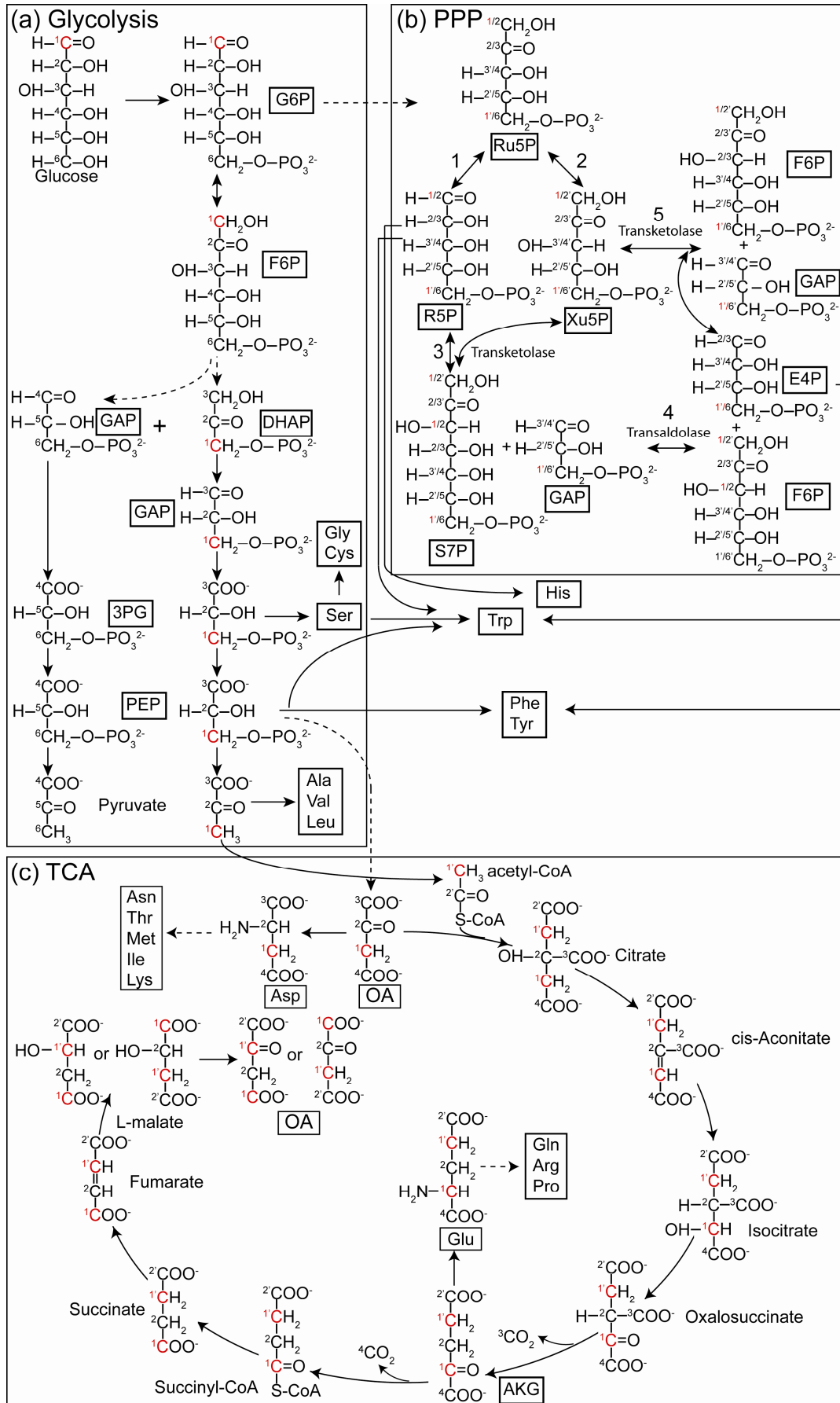


Figure A8. Biosynthetic pathways of group III amino acids for proteins expressed in an *E.coli* medium containing [2-¹³C]Glc as the sole carbon source. (a) OA derived amino acids and (c) AKG derived amino acids. The carbons that are ¹³C labeled are colored in red. The amino acids derived from OA and AKG have multiple labeling patterns. For clarity, we show the biosynthetic pathways of these amino acids, and list all the types of (b) OA and (d) AKG. Abbreviations: L, L- α,ϵ -diaminopimelate, DAP.



Appendix A

Figure A9. Key metabolic pathways involved in the biosynthesis of 20 amino acid when *E. coli* is grown on [1-¹³C]Glc as the sole carbon source, i.e., (a) glycolysis, (b) PPP, and (c) TCA cycle. The *superscript numbers* do not correspond to the nomenclature number of a given molecule but indicate the original carbon position in the six-carbon [1-¹³C]Glc chain. The carbons that are ¹³C labeled are colored in red. For the sake of clarity, sugars are shown in their linear forms. For simplicity, some steps are not included, which are indicated by dashed arrows. Double-headed arrows indicate reversible reactions.

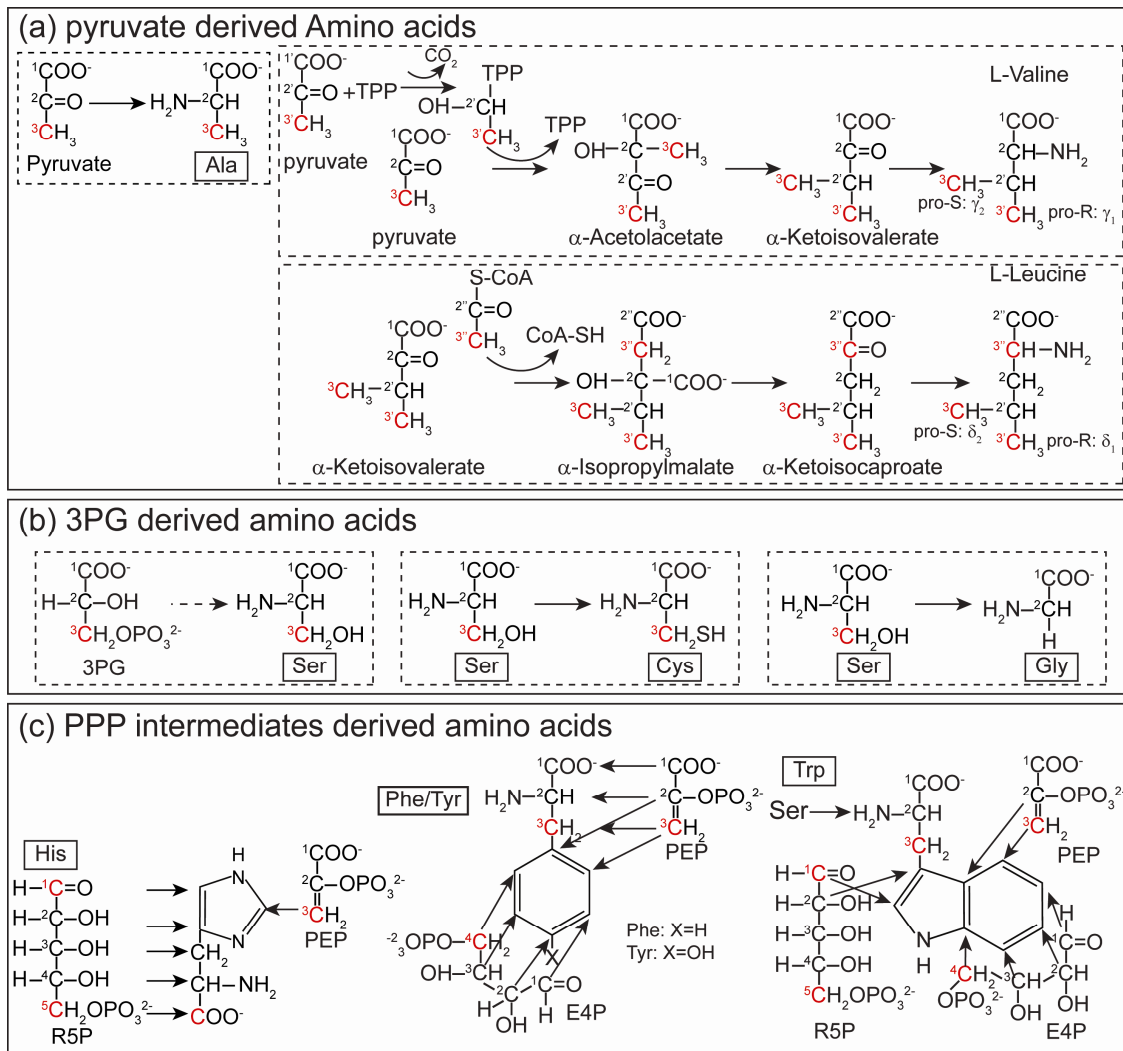


Figure A10. Biosynthetic pathways of the group I and II amino acids for proteins expressed in an *E. coli* medium containing [1-¹³C]Glc as the sole carbon source. The carbons that are ¹³C labeled are colored in red.

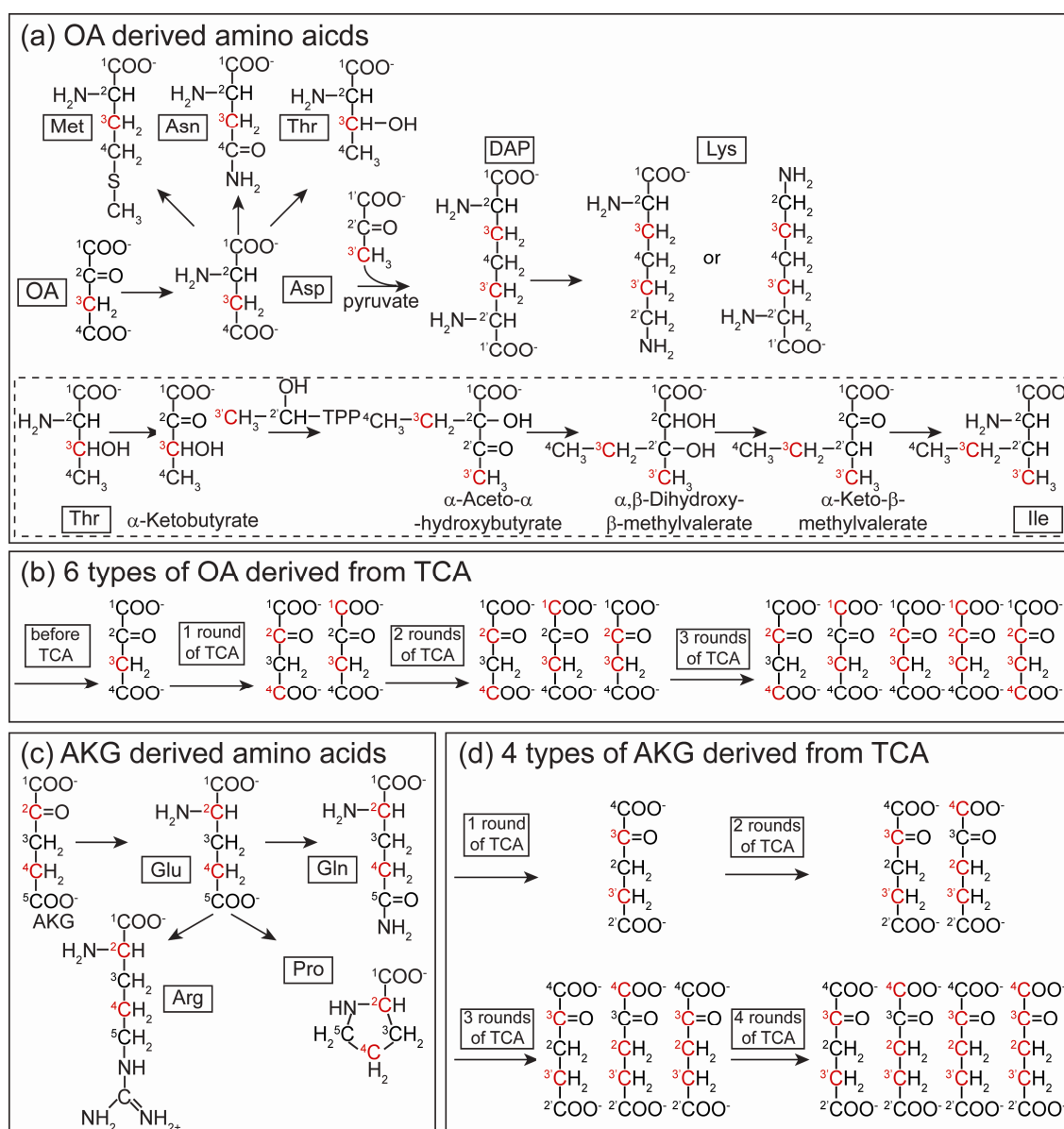


Figure A11. Biosynthetic pathways of the group III amino acids for proteins expressed in an *E.coli* medium containing [1-¹³C]Glc as the sole carbon source. All the isotopomers of OA and AKG derived from the TCA cycle are summarized in (b) and (d), respectively. The carbons that are ¹³C labeled are colored in red.

Table A1. Details of ssNMR experiments used for sequential resonance assignment.

Samples	Dimensional	Rotor [mm]	Experiment	¹ H freq. [MHz]	MAS freq. [kHz]	Mixing time [ms]	CP contact time [μs]	TD ₁	aq ₁ [ms]	Number of scans	Delay time [s]	Total measurement time
[U- ¹³ C]Glc-labeled m.c.S sample	D	2	DREAM	800	18	4	700/ -	480	11.7	176	2	2 d
		4	PDS	800	11	20	900/-/-	768	8.7	96	2	1 d 18 h
		4	PDS	850	11	20	600/ -/ -	640	6.8	176	2	2 d 16 h
		4	PDS	850	11	80	600/ -/ -	768	8.2	128	2	2 d 9 h
		4	PDS	850	11	150	600/ -/ -	768	8.2	192	2	3 d 16 h
		4	PDS	800	11	250	700/ -/ -	768	8.7	288	2	5 d 19 h
		4	PDS	600	11	100	1200/ -/ / -	640	8.8	192	2	3 d
		4	NCA	800	11		700/ -/ -	64	8	1152	2	1 d 17 h
		4	NCO	800	11		700/ -/ -	64	8	640	2	23 h
		4	NCACX	800	11	60 (PDS)	800/ 3000/ -	68	8.2	2976	2	4 d 21 h
		4	NCOCX	800	11	50 (PDS)	800/ 3000/ -	68	8.2	2976	2	4 d 20 h
		4	NCOCX	800	11	25 (PDS)	600/ 3000/ -	68	8.2	2784	2	4 d 11 h
		3.2	NCACX	800	15	50 (DARR)	1000/ 3000/ -	384	11.3	640	2	5 d 22 h
		3.2	NCOCX	800	15	50 (DARR)	900/ 3000/ -	384	11.3	672	2	6 d 10 h
		4	NN_PDS D	800	11	6000	900/ -/ -	160	11.5	192	2	2 d 20 h
		4	NN_PDS D	800	11	10 000	900/ -/ -	160	11.5	144	2	3 d 5 h

	4	(2Q,1Q) correlation	800	8	0.5	700/-/ -	312	5.9	384	2	2 d 19 h	
		4	NHHC	850	11	0.15	600/ 200/20 0	70	8.1	1792	2	2 d 22h
	3 D	4	NCACX	850	11	50	600/ 3000/-	48	5.4	96	2.3	5 d 7 h
		4	NCOCX	850	11	50	600/ 3000/-	48	5.4	96	2.5	6 d 14 h
[2- ¹³ C]Glc-labeled mαS sample	2 D	4	PDS D	850	11	500	400/-/ -	1411	15	160	2.5	7 d 11 h
		4	PDS D	800	11	300	800/-/ -	900	12	304	2	7 d 9 h
		4	PDS D	800	11	100	700/-/ -	564	14	224	2.5	3 d 20 h
		4	PDS D	800	11	700	700/-/ -	1104	12.5	272	2.4	7 d 21 h
		4	PDS D	800	11	1000	700/-/ -	546	13.6	304	2.5	6 d 20 h
		4	PDS D	600	11	100	800/-/ -	680	10.7	128	2.5	2 d 20 h
		4	PDS D	600	11	400	800/-/ -	660	10.4	240	2.5	5 d 15 h
		4	PDS D	600	11	1000	700/-/ -	556	8	576	2	11 d 16 h
		4	NCA	850	11		800/ 3000	160	14.5	576	2.5	2 d 17 h
		4	NN_SD	850	11	6000	400/-/ -	218	14.9	128	2.5	2 d 21 h
	4	NHHC	850	11	0.175	400/ 400/20 0	128	14.9	1376	2.4	4 d 18 h	
	4	NCO	850	11		800/30 00	128	14.9	736	2.4	2 d 16 h	
3 D	4	NCACX	850	11	800	600/ 1000/-	30	5.1	224	2.3	10 d 1 h	

Appendix A

[1- ¹³ C]Glc-labeled mαS sample	2 D	4	PDS	850	11	100	600/- -	510	10	400	2.5	6 d 5 h
		4	PDS	800	11	400	1000/- /-	640	13.1	432	3.5	12 d 15 h
		4	PDS	850	11	1000	700	800	8.5	464	2.1	13 d 11 h
		4	NHHC	850	11	0.6	800/ 400/20 0	62	9	2464	2.1	3 d 19 h
		4	NCA	850	11		800/ 3000	70	11.9	816	3	2 d
		4	NCO	850	11		800/ 3000	58	9.9	1072	2.6	1 d 21 h
[M- ¹³ C/ ¹⁵ N]-labeled mαS sample	2 D	4	NHHC	850	11	0.5	700/ 400/20 0	54	7.8	7648	2.2	10 d 14 h

Table A2. Chemical shift assignments for m α S fibrils.

Res.	N	C'	C α	C β	C γ 2/1	C δ	C ϵ	C ζ	Nsc
Ser?	-	-	58.3	63.9					
Val40*	-	-	60.4	35.2	-/-				
Gly41	119	174.7	47.9						
Ser42	112.2	171.5	58.7	67.6					
Lys43	122.3	175.7	54.6	35.8	25.6	30.2	42.1		-
Thr44	113.4	175.2	59.5	71.4	22.4				
Lys45	123.3	173.5	56.2	36.8	26.8	-	42.2		-
Glu46	126.5	174.6	53.9	32.9	35.6	183.1			
Gly47	115.1	172.6	48.4						
Val48	119.2	174.1	59.8	37.7	24.2/21.2				
Val49	126.3	174.6	60.8	34.6	23.1/-				
His50	-	-	-	-	-	-	-		-/-
Gly51	-	174.9	48.6						
Val52	124.7	175.3	61.1	33.5	20.7/-				
Thr53	127.8	173.1	61.6	70.4	21.2				
Thr54	127.9	172.7	61.9	70.4	21.2				
Val55	127.5	174.2	61	35.7	21.2/-				
Ala56	131.6	175.3	50.5	22.3					
Glu57	122.5	175.9	53.7	33.6	35.7	183.3			
Lys58	118.1	175	57.9	30.1	27.6	31.9	42.5		34.1
Thr59	106.5	175	61.1	69.8	24.2				
Lys60	122.1	175.3	56.1	36.3	26.1	30.1	42.3		29.9
Glu61	128.4	174.3	55	33.7	-	-			
Gln62	130	174	54.7	32.5	34.2	179.9			111.9
Val63	124.9	175	61	36.3	21/22.2				
Thr64	126.9	172.7	62.4	69.8	21.5				
Asn65	125.1	172.7	51.7	43.2	175				115.2
Val66	127.2	178.5	60.7	33.7	19.9/21.1				
Gly67	111.2	172.8	46.7						
Gly68	103.4	172.2	43.3						
Ala69	126.9	175.3	50.4	23.4					
Val70	120.8	174.4	60.3	36.1	21.9/-				
Val71	127.3	176.5	60.8	35.3	21.7/20.7				
Thr72	114.7	175.9	60.3	69.4	21.7				
Gly73	110.6	173.7	44.3						
Val74	124.9	175.3	61.5	35	21.3/19.2				

Appendix A

Thr75	127.9	172	61.7	70.1	21.1			
Ala76	130.5	174.1	49.6	21.2				
Val77	124.1	172.7	60.6	35.8	20.3/21.5			
Ala78	130.5	176.2	49.8	24.7				
Gln79	120.6	176.4	52.4	32.6	32.9	177.7		111.2
Lys80	123.8	176.2	60.5	31.4	27.3	29.2	42.1	33.2
Thr81	115	173.6	61.8	72	22.7			
Val82	126.3	174.2	61.1	34.4	20.5/20.1			
Glu83	123.4	173	54.9	29.5	34.9	183.9		
Gly84	106.7	174.1	44.8					
(**)		(173.2)	(44.6)					
Ala85	126.6	177.8	50.3	21.9				
(**)	(125.8)		(49.7)					
Gly86	110.3	174.7	47.4					
(**)			(47.3)					
Asn87	117.8	175.9	53.6	39.9	172.9			116.6
Ile88	122.4	175.6	60.5	39.7	17.7/27.3	13.4		
(**)						(13.1)		
Ala89	129.5	176.6	54.9	19.3				
(**)			(55.4)					
Ala90	123.2	174.6	51.4	21.2				
Ala91	127.8	175.5	49.6	21.9				
Thr92	125.2	174.6	60.9	70.2	21.8			
Gly93	115.3	170.4	47.7					
(**)			(47.5)					
Phe94	125.9	-	55.9	45.1	136.4	132.1	130.9	-
(***)	(126)	(173.5)	(54.6)	(45.5)	(138.1)	(132.2)	(129.3)	(-)
Val95	128.5	-	60.5	35.9	22.4/-			
Asp?	-	-	54.9	42.9	-			
Pro?	-	-	62.9	32.1	27.3	50.7		

([∞]) The assignment of Val40 is ambiguous.

(**) Resonances in brackets indicate peak doubling.

(***) Phe94 has a second set of resonances (Phe94').

Appendix B

Stereospecific Assignment of Val and Leu Prochiral Methyl Groups

Appendix B

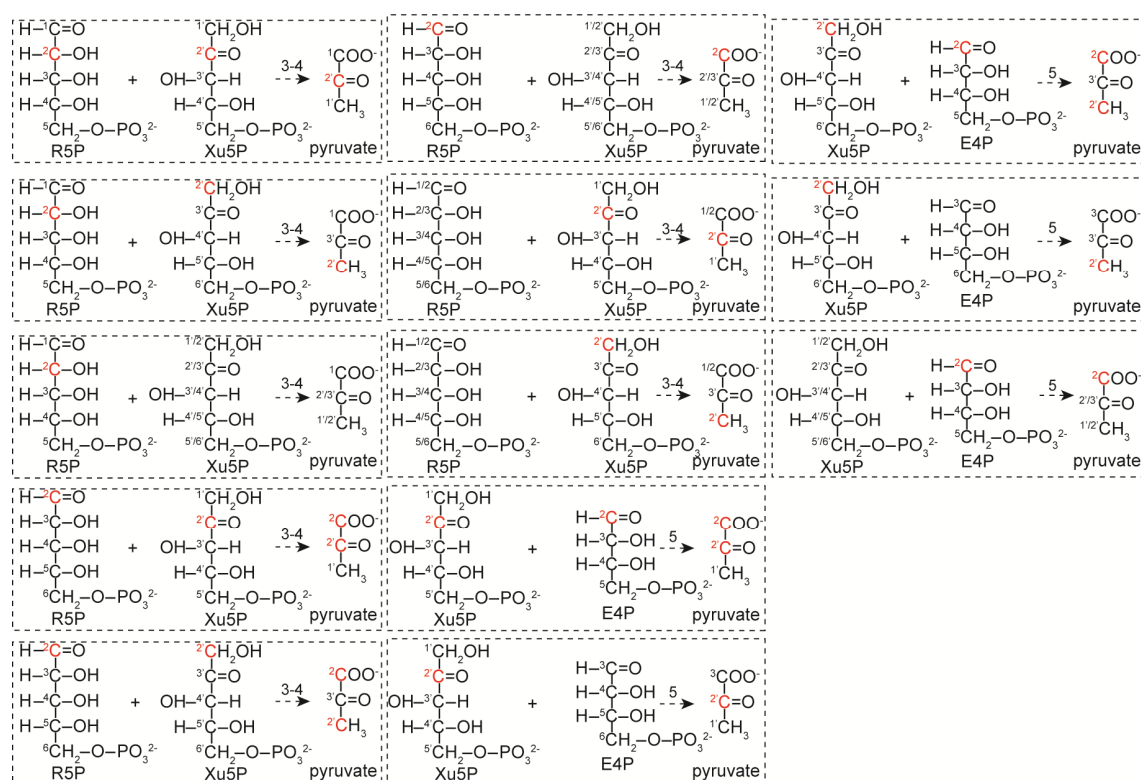


Figure B1. Summary of the isotopomers of pyruvate formed in PPP. In total, 6 isotopomers of pyruvate are obtained, [2-¹³C]pyruvate, [3-¹³C]pyruvate, [1,3-¹³C]pyruvate, [1,2-¹³C]pyruvate, [1-¹³C]pyruvate, and [1,2,3-¹²C]pyruvate.

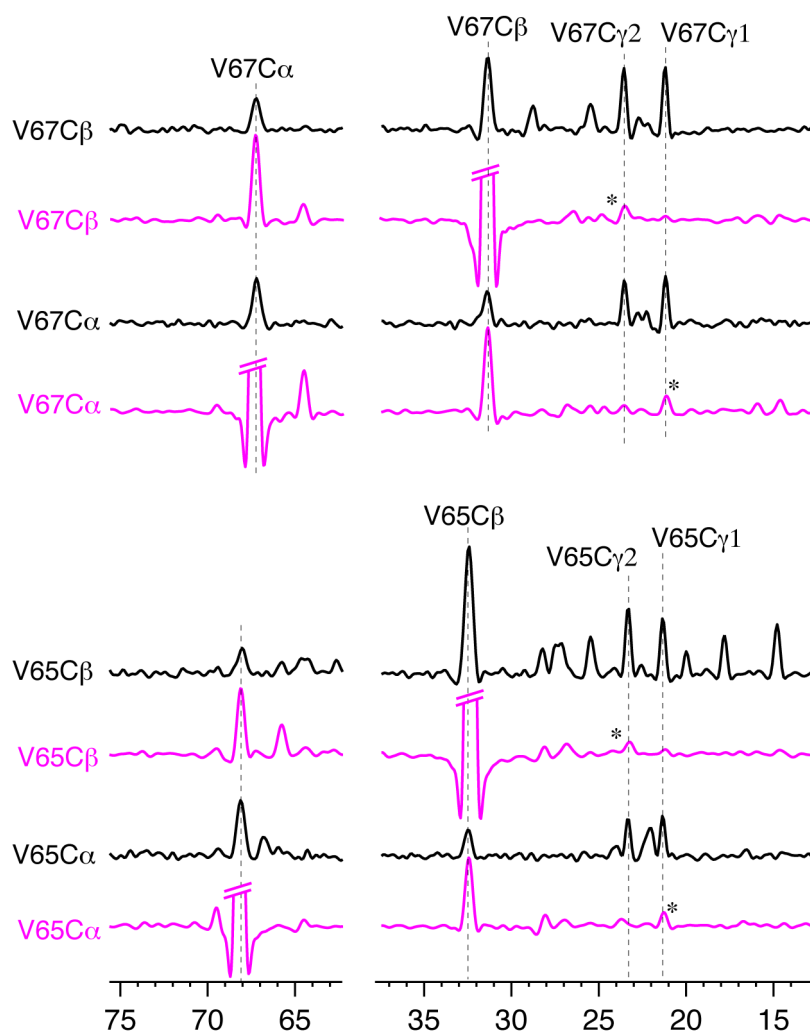


Figure B2. 1D slices for the residues Val65 and Val67 in the 2D ^{13}C - ^{13}C PDSD spectra of $[\text{U-}^{13}\text{C}]\text{Glc}$ - (in black) and $[\text{2-}^{13}\text{C}]\text{Glc}$ -labeled (in magenta) T3SS PrgI needles shown in Fig. 3.4.

References

- [1] R.G. Griffin, Dipolar recoupling in MAS spectra of biological solids. *Nat. Struct. Biol.* 5 Suppl (1998) 508-512.
- [2] A. McDermott, Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. *Annu. Rev. Biophys.* 38 (2009) 385-403.
- [3] M. Renault, A. Cukkemane, and M. Baldus, Solid-state NMR spectroscopy on complex biomolecules. *Angew. Chem. Int. Ed. Engl.* 49 (2010) 8346-8357.
- [4] A. Bockmann, 3D protein structures by solid-state NMR spectroscopy: ready for high resolution. *Angew. Chem. Int. Ed. Engl.* 47 (2008) 6110-6113.
- [5] M. Hong, Determination of Multiple ϕ -Torsion Angles in Proteins by Selective and Extensive ^{13}C Labeling and Two-Dimensional Solid-State NMR. *J. Magn. Reson.* 139 (1999) 389-401.
- [6] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, and H. Oschkinat, Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420 (2002) 98-102.
- [7] C.G. Hoogstraten, and J.E. Johnson, Metabolic labeling: Taking advantage of bacterial pathways to prepare spectroscopically useful isotope patterns in proteins and nucleic acids. *Concept. Magn. Reson. A.* 32A (2008) 34-55.
- [8] A. Loquet, B. Bardiaux, C. Gardiennet, C. Blanchet, M. Baldus, M. Nilges, T. Malliavin, and A. Bockmann, 3D structure determination of the Crh protein from highly ambiguous solid-state NMR restraints. *J. Am. Chem. Soc.* 130 (2008) 3579-3589.
- [9] T. Manolikas, T. Herrmann, and B.H. Meier, Protein structure determination from ^{13}C spin-diffusion solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 130 (2008) 3959-3966.
- [10] S.D. Cady, K. Schmidt-Rohr, J. Wang, C.S. Soto, W.F. Degrado, and M. Hong, Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* 463 (2010) 689-692.
- [11] L. Shi, M.A. Ahmed, W. Zhang, G. Whited, L.S. Brown, and V. Ladizhansky, Three-dimensional solid-state NMR study of a seven-helical integral membrane proton pump--structural insights. *J. Mol. Biol.* 386 (2009) 1078-1093.
- [12] R. Schneider, M. Etzkorn, K. Giller, V. Daebel, J. Einfeld, M. Zweckstetter, C. Griesinger, S. Becker, and A. Lange, The Native Conformation of the Human VDAC1 N Terminus. *Angew. Chem. Int. Ed. Engl.* 49 (2010) 1882-1885.
- [13] A. Lange, K. Giller, S. Hornig, M.F. Martin-Eauclaire, O. Pongs, S. Becker, and M. Baldus, Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature* 440 (2006) 959-962.
- [14] J. Yang, L. Aslimovska, and C. Glaubitz, Molecular dynamics of proteorhodopsin in lipid bilayers by solid-state NMR. *J. Am. Chem. Soc.* 133 (2011) 4874-4881.
- [15] T. Jacso, W.T. Franks, H. Rose, U. Fink, J. Broecker, S. Keller, H. Oschkinat, and B. Reif, Characterization of Membrane Proteins in Isolated Native Cellular

- Membranes by Dynamic Nuclear Polarization Solid-State NMR Spectroscopy without Purification and Reconstitution. *Angew. Chem. Int. Ed. Engl.* (2011).
- [16] S. Chimon, M.A. Shaibat, C.R. Jones, D.C. Calero, B. Aizezi, and Y. Ishii, Evidence of fibril-like beta-sheet structures in a neurotoxic amyloid intermediate of Alzheimer's beta-amyloid. *Nat. Struct. Mol. Biol.* 14 (2007) 1157-1164.
- [17] S. Jehle, P. Rajagopal, B. Bardiaux, S. Markovic, R. Kuhne, J.R. Stout, V.A. Higman, R.E. Klevit, B.J. van Rossum, and H. Oschkinat, Solid-state NMR and SAXS studies provide a structural basis for the activation of alphaB-crystallin oligomers. *Nat. Struct. Mol. Biol.* 17 (2010) 1037-1042.
- [18] Y. Han, J. Ahn, J. Concel, I.J. Byeon, A.M. Gronenborn, J. Yang, and T. Polenova, Solid-state NMR studies of HIV-1 capsid protein assemblies. *J. Am. Chem. Soc.* 132 (2010) 1976-1987.
- [19] B. Chen, and R. Tycko, Structural and dynamical characterization of tubular HIV-1 capsid protein assemblies by solid state nuclear magnetic resonance and electron microscopy. *Protein Sci.* 19 (2010) 716-730.
- [20] A. Goldbourn, L.A. Day, and A.E. McDermott, Intersubunit hydrophobic interactions in Pf1 filamentous phage. *J. Biol. Chem.* 285 (2010) 37051-37059.
- [21] O. Morag, G. Abramov, and A. Goldbourn, Similarities and Differences Within Members of the Ff Family of Filamentous Bacteriophage Viruses. *J. Phys. Chem. B* (2011).
- [22] C.P. Jaroniec, C.E. MacPhee, V.S. Bajaj, M.T. McMahon, C.M. Dobson, and R.G. Griffin, High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* 101 (2004) 711-716.
- [23] J.J. Helmus, K. Surewicz, P.S. Nadaud, W.K. Surewicz, and C.P. Jaroniec, Molecular conformation and dynamics of the Y145Stop variant of human prion protein in amyloid fibrils. *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 6284-6289.
- [24] I. Bertini, L. Gonnelli, C. Luchinat, J. Mao, and A. Nesi, A new structural model of Abeta40 fibrils. *J. Am. Chem. Soc.* 133 (2011) 16013-16022.
- [25] A.T. Petkova, W.M. Yau, and R. Tycko, Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry* 45 (2006) 498-512.
- [26] J.T. Nielsen, M. Bjerring, M.D. Jeppesen, R.O. Pedersen, J.M. Pedersen, K.L. Hein, T. Vosegaard, T. Skrydstrup, D.E. Otzen, and N.C. Nielsen, Unique identification of supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* 48 (2009) 2118-2121.
- [27] A. Loquet, L. Bousset, C. Gardiennet, Y. Sourigues, C. Wasmer, B. Habenstein, A. Schutz, B.H. Meier, R. Melki, and A. Bockmann, Prion fibrils of Ure2p assembled under physiological conditions contain highly ordered, natively folded modules. *J. Mol. Biol.* 394 (2009) 108-118.
- [28] J.J. Helmus, K. Surewicz, W.K. Surewicz, and C.P. Jaroniec, Conformational flexibility of Y145Stop human prion protein amyloid fibrils probed by solid-state nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* 132 (2010) 2393-2403.

References

- [29] C. Wasmer, A. Lange, H. Van Melckebeke, A.B. Siemer, R. Riek, and B.H. Meier, Amyloid Fibrils of the HET-s(218-289) Prion Form a β Solenoid with a Triangular Hydrophobic Core. *Science* 319 (2008) 1523-1526.
- [30] D.D. Laws, H.M. Bitter, and A. Jerschow, Solid-State NMR Spectroscopic Methods in Chemistry. *Angew. Chem. Int. Ed. Engl.* 41 (2002) 3096-3129.
- [31] J. Cavanagh, W.J. Fairbrother, A.G.P. III, M. Rance, and N.J. Skelton, *Protein NMR spectroscopy: principles and practice.*, Elsevier: Academic press, 2007.
- [32] M. Sattler, J. Schleucher, and C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* 34 (1999) 93-158.
- [33] E.R. Andrew, A. Bradbury, and R.G. Eades, Nuclear Magnetic Resonance Spectra from a Crystal Rotated at High Speed. *Nature* 182 (1958) 1659-1659.
- [34] I.J. Lowe, Free Induction Decays of Rotating Solids. *Phys. Rev. Lett.* 2 (1959) 285-287.
- [35] A.E. Bennett, C.M. Rienstra, M. Auger, K.V. Lakshmi, and R.G. Griffin, Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* 103 (1995) 6951-6958.
- [36] B.M. Fung, A.K. Khitrin, and K. Ermolaev, An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.* 142 (2000) 97-101.
- [37] M. Baldus, Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning. *Prog. Nucl. Magn. Reson. Spectrosc.* 41 (2002) 1-47.
- [38] J.M. Griffiths, and R.G. Griffin, Nuclear magnetic resonance methods for measuring dipolar couplings in rotating solids. *Anal. Chim. Acta.* 283 (1993) 1081-1101.
- [39] A. Pines, M.G. Gibby, and J.S. Waugh, Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* 59 (1973) 569-590.
- [40] S.R. Hartmann, and E.L. Hahn, Nuclear Double Resonance in the Rotating Frame. *Phys. Rev.* 128 (1962) 2042-2053.
- [41] G. Metz, X.L. Wu, and S.O. Smith, Ramped-Amplitude Cross-Polarization in Magic-Angle-Spinning Nmr. *J. Magn. Reson. A* 110 (1994) 219-227.
- [42] S. Hediger, B.H. Meier, and R.R. Ernst, Adiabatic Passage Hartmann-Hahn Cross-Polarization in Nmr under Magic-Angle Sample-Spinning. *Chem. Phys. Lett.* 240 (1995) 449-456.
- [43] O.C. Andronesi, S. Becker, K. Seidel, H. Heise, H.S. Young, and M. Baldus, Determination of Membrane Protein Structure and Dynamics by Magic-Angle-Spinning Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 127 (2005) 12965-12974.
- [44] H. Heise, W. Hoyer, S. Becker, O.C. Andronesi, D. Riedel, and M. Baldus, Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* 102 (2005) 15871-15876.
- [45] A. Kubo, and C.A. McDowell, Spectral Spin Diffusion in Polycrystalline Solids under Magic-angle Spinning. *J. Chem. Soc., Faraday Trans. 1* 84 (1988) 3713-3730.

- [46] D. Suter, and R.R. Ernst, Spin diffusion in resolved solid-state NMR spectra. *Phys. Rev. B. Condens. Matter.* 32 (1985) 5608-5627.
- [47] D. Suter, and R.R. Ernst, Spectral spin diffusion in the presence of an extraneous dipolar reservoir. *Phys. Rev. B* 25 (1982) 6038-6041.
- [48] K. Takegoshi, S. Nakamura, and T. Terao, C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* 344 (2001) 631-637.
- [49] M. Baldus, A.T. Petkova, J. Herzfeld, and R.G. Griffin, Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* 95 (1998) 1197-1207.
- [50] A.T. Petkova, M. Baldus, M. Belenky, M. Hong, R.G. Griffin, and J. Herzfeld, Backbone and side chain assignment strategies for multiply labeled membrane peptides and proteins in the solid state. *J. Magn. Reson.* 160 (2003) 1-12.
- [51] A. Lange, S. Luca, and M. Baldus, Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* 124 (2002) 9704-9705.
- [52] P. Hodgkinson, and L. Emsley, The accuracy of distance measurements in solid-state NMR. *J. Magn. Reson.* 139 (1999) 46-59.
- [53] S. Kiihne, M.A. Mehta, J.A. Stringer, D.M. Gregory, J.C. Shiels, and G.P. Drobny, Distance Measurements by Dipolar Recoupling Two-Dimensional Solid-State NMR. *J. Phys. Chem. A.* 102 (1998) 2274-2282.
- [54] L.-Y. Lian, and D.A. Middleton, Labelling approaches for protein structural studies by solution-state and solid-state NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 39 (2001) 171-190.
- [55] G.W. Vuister, S.-J. Kim, C. Wu, and A. Bax, 2D and 3D NMR Study of Phenylalanine Residues in Proteins by Reverse Isotopic Labeling. *J. Am. Chem. Soc.* 116 (1994) 9206-9210.
- [56] M. Etzkorn, S. Martell, O.C. Andronesi, K. Seidel, M. Engelhard, and M. Baldus, Secondary Structure, Dynamics, and Topology of a Seven-Helix Receptor in Native Membranes, Studied by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed. Engl.* 46 (2007) 459-462.
- [57] L. Skrisovska, M. Schubert, and F.H.-T. Allain, Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins. *J. Biomol. NMR.* 46 (2010) 51-65.
- [58] M. Hong, Determination of multiple phi-torsion angles in proteins by selective and extensive C-13 labeling and two-dimensional solid-state NMR. *J. Magn. Reson.* 139 (1999) 389-401.
- [59] D.M. LeMaster, and D.M. Kushlan, Dynamical Mapping of *E. coli* Thioredoxin via ¹³C NMR Relaxation Analysis. *J. Am. Chem. Soc.* 118 (1996) 9255-9264.
- [60] M. Hong, and K. Jakes, Selective and extensive ¹³C labeling of a membrane protein for solid-state NMR investigations. *J. Biomol. NMR.* 14 (1999) 71-74.
- [61] P. Lundström, K. Teilum, T. Carstensen, I. Bezsonova, S. Wiesner, D.F. Hansen, T.L. Religa, M. Akke, and L.E. Kay, Fractional ¹³C enrichment of isolated carbons using [1-¹³C]- or [2-¹³C]-glucose facilitates the accurate measurement of

References

- dynamics at backbone C^α and side-chain methyl positions in proteins. *J. Biomol. NMR.* 38 (2007) 199-212.
- [62] V.A. Higman, J. Flinders, M. Hiller, S. Jehle, S. Markovic, S. Fiedler, B.J. van Rossum, and H. Oschkinat, Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively ¹³C-labelled proteins. *J. Biomol. NMR.* 44 (2009) 245-60.
- [63] J.E. Galvin, V.M. Lee, and J.Q. Trojanowski, Synucleinopathies: Clinical and Pathological Implications. *Arch. Neurol.* 58 (2001) 186-190.
- [64] M. Goedert, α -Synuclein and Neurodegenerative Diseases. *Nat. Rev. Neurosci.* 2 (2001) 492-501.
- [65] M.G. Spillantini, and M. Goedert, The α -Synucleinopathies: Parkinson's Disease, Dementia with Lewy Bodies, and Multiple System Atrophy. *Ann. N. Y. Acad. Sci.* 920 (2000) 16-27.
- [66] L.S. Forno, Neuropathology of Parkinson's disease. *J. Neuropathol. Exp. Neurol.* 55 (1996) 259-272.
- [67] M.G. Spillantini, M.L. Schmidt, V.M.Y. Lee, J.Q. Trojanowski, R. Jakes, and M. Goedert, alpha-synuclein in Lewy bodies. *Nature* 388 (1997) 839-840.
- [68] M.G. Spillantini, R.A. Crowther, R. Jakes, M. Hasegawa, and M. Goedert, alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies. *Proc. Natl. Acad. Sci. U.S.A.* 95 (1998) 6469-6473.
- [69] M. Baba, S. Nakajo, P.H. Tu, T. Tomita, K. Nakaya, V.M.Y. Lee, J.Q. Trojanowski, and T. Iwatsubo, Aggregation of alpha-synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with lewy bodies. *Am. J. Pathol.* 152 (1998) 879-884.
- [70] P.H. Weinreb, W.G. Zhen, A.W. Poon, K.A. Conway, and P.T. Lansbury, NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* 35 (1996) 13709-13715.
- [71] D. Eliezer, E. Kutluay, R. Bussell, Jr., and G. Browne, Conformational properties of alpha-synuclein in its free and lipid-associated states. *J. Mol. Biol.* 307 (2001) 1061-1073.
- [72] W.S. Davidson, A. Jonas, D.F. Clayton, and J.M. George, Stabilization of alpha-synuclein secondary structure upon binding to synthetic membranes. *J. Biol. Chem.* 273 (1998) 9443-9449.
- [73] B. Caughey, and P.T. Lansbury, PROTOFIBRILS, PORES, FIBRILS, AND NEURODEGENERATION: Separating the Responsible Protein Aggregates from The Innocent Bystanders. *Annu. Rev. Neurosci.* 26 (2003) 267-298.
- [74] T. Bartels, J.G. Choi, and D.J. Selkoe, alpha-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* 477 (2011) 107-110.
- [75] W. Wang, I. Perovic, J. Chittuluru, A. Kaganovich, L.T. Nguyen, J. Liao, J.R. Auclair, D. Johnson, A. Landeru, A.K. Simorellis, S. Ju, M.R. Cookson, F.J. Asturias, J.N. Agar, B.N. Webb, C. Kang, D. Ringe, G.A. Petsko, T.C. Pochapsky, and Q.Q. Hoang, A soluble {alpha}-synuclein construct forms a dynamic tetramer. *Proc. Natl. Acad. Sci. U.S.A.* (2011).

- [76] K.A. Conway, J.D. Harper, and P.T. Lansbury, Fibrils formed in vitro from alpha-synuclein and two mutant forms linked to Parkinson's disease are typical amyloid. *Biochemistry* 39 (2000) 2552-2563.
- [77] L.C. Serpell, J. Berriman, R. Jakes, M. Goedert, and R.A. Crowther, Fiber diffraction of synthetic alpha-synuclein filaments shows amyloid-like cross-beta conformation. *Proc. Natl. Acad. Sci. U.S.A.* 97 (2000) 4897-4902.
- [78] H. Miake, H. Mizusawa, T. Iwatsubo, and M. Hasegawa, Biochemical characterization of the core structure of alpha-synuclein filaments. *J. Biol. Chem.* 277 (2002) 19213-19219.
- [79] K.D. Kloepper, D.H. Zhou, Y. Li, K.A. Winter, J.M. George, and C.M. Rienstra, Temperature-dependent sensitivity enhancement of solid-state NMR spectra of alpha-synuclein fibrils. *J. Biomol. NMR* 39 (2007) 197-211.
- [80] M. Vilar, H.T. Chou, T. Luhrs, S.K. Maji, D. Riek-Loher, R. Verel, G. Manning, H. Stahlberg, and R. Riek, The fold of alpha-synuclein fibrils. *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 8637-8642.
- [81] G. Comellas, L.R. Lemkau, A.J. Nieuwkoop, K.D. Kloepper, D.T. Lador, R. Ebisu, W.S. Woods, A.S. Lipton, J.M. George, and C.M. Rienstra, Structured Regions of alpha-Synuclein Fibrils Include the Early-Onset Parkinson's Disease Mutation Sites. *J. Mol. Biol.* 411 (2011) 881-895.
- [82] J. Gath, B. Habenstein, L. Bousset, R. Melki, B.H. Meier, and A. Bockmann, Solid-state NMR sequential assignments of alpha-synuclein. *Biomol. NMR. Assign.* (2011).
- [83] C. Del Mar, E.A. Greenbaum, L. Mayne, S.W. Englander, and V.L. Woods, Jr., Structure and properties of α -synuclein and other amyloids determined at the amino acid level. *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 15477-15482.
- [84] H. Heise, M.S. Celej, S. Becker, D. Riede, A. Pelah, A. Kumar, T.M. Jovin, and M. Baldus, Solid-state NMR reveals structural differences between fibrils of wild-type and disease-related A53T mutant alpha-synuclein. *J. Mol. Biol.* 380 (2008) 444-450.
- [85] L. Kang, K.P. Wu, M. Vendruscolo, and J. Baum, The A53T Mutation is Key in Defining the Differences in the Aggregation Kinetics of Human and Mouse alpha-Synuclein. *J. Am. Chem. Soc.* 133 (2011) 13465-13470.
- [86] L. Hong, H.W. Ko, B.J. Gwag, E. Joe, S. Lee, Y.T. Kim, and Y.H. Suh, The cDNA cloning and ontogeny of mouse alpha-synuclein. *Neuroreport* 9 (1998) 1239-1243.
- [87] J.C. Rochet, K.A. Conway, and P.T. Lansbury, Inhibition of fibrillization and accumulation of prefibrillar oligomers in mixtures of human and mouse alpha-synuclein. *Biochemistry* 39 (2000) 10619-10626.
- [88] K.C. Luk, V. Kehm, J. Carroll, B. Zhang, P. O'Brien, J.Q. Trojanowski, and V.M. Lee, Pathological α -Synuclein Transmission Initiates Parkinson-like Neurodegeneration in Nontransgenic Mice. *Science* 338 (2012) 949-953.
- [89] A. Loquet, G. Lv, K. Giller, S. Becker, and A. Lange, ^{13}C spin dilution for simplified and complete solid-state NMR resonance assignment of insoluble biological assemblies. *J. Am. Chem. Soc.* 133 (2011) 4722-4725.

References

- [90] M. Etzkorn, A. Bockmann, A. Lange, and M. Baldus, Probing molecular interfaces using 2D magic-angle-spinning NMR on protein mixtures with different uniform labeling. *J. Am. Chem. Soc.* 126 (2004) 14746-14751.
- [91] A. Loquet, K. Giller, S. Becker, and A. Lange, Supramolecular Interactions Probed by ^{13}C - ^{13}C Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 132 (2010) 15164-15166.
- [92] W. Hoyer, T. Antony, D. Cherny, G. Heim, T.M. Jovin, and V. Subramaniam, Dependence of α -Synuclein Aggregate Morphology on Solution Conditions. *J. Mol. Biol.* 322 (2002) 383-393.
- [93] W. Hoyer, D. Cherny, V. Subramaniam, and T.M. Jovin, Impact of the acidic C-terminal region comprising amino acids 109-140 on alpha-synuclein aggregation in vitro. *Biochemistry* 43 (2004) 16233-16242.
- [94] J.L. Markley, A. Bax, Y. Arata, C.W. Hilbers, R. Kaptein, B.D. Sykes, P.E. Wright, and K. Wuthrich, Recommendations for the presentation of NMR structures of proteins and nucleic acids--IUPAC-IUBMB-IUPAB Inter-Union Task Group on the standardization of data bases of protein and nucleic acid structures determined by NMR spectroscopy. *Eur. J. Biochem.* 256 (1998) 1-15.
- [95] C.R. Morcombe, and K.W. Zilm, Chemical shift referencing in MAS solid state NMR. *J. Magn. Reson.* 162 (2003) 479-486.
- [96] K. Schmidt-Rohr, and H.W. Spiess, *Multidimensional Solid-State NMR and Polymers*, Academic Press; 1 edition (November 29, 1994), 1994.
- [97] R. Verel, M. Baldus, M. Ernst, and B.H. Meier, A homonuclear spin-pair filter for solid-state NMR based on adiabatic-passage techniques. *Chem. Phys. Lett.* 287 (1998) 421-428.
- [98] M. Hohwy, C.M. Rienstra, C.P. Jaroniec, and R.G. Griffin, Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. *J. Chem. Phys.* 110 (1999) 7983-7992.
- [99] B. Reif, M. Hohwy, C.P. Jaroniec, C.M. Rienstra, and R.G. Griffin, NH-NH Vector Correlation in Peptides by Solid-State NMR. *J. Magn. Reson.* 145 (2000) 132-141.
- [100] D. Voet, and J.G. Voet, *Biochemistry*, John Wiley & Sons, Inc., 2011.
- [101] J. Katz, and R. Rognstad, The labeling of pentose phosphate from glucose- ^{14}C and estimation of the rates of transaldolase, transketolase, the contribution of the pentose cycle, and ribose phosphate synthesis. *Biochemistry* 6 (1967) 2227-47.
- [102] K. Teilum, U. Brath, P. Lundström, and M. Akke, Biosynthetic ^{13}C Labeling of Aromatic Side Chains in Proteins for NMR Relaxation Measurements. *J. Am. Chem. Soc.* 128 (2006) 2506-2507.
- [103] T. Szyperski, Biosynthetically directed fractional ^{13}C -labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *Eur. J. Biochem.* 232 (1995) 433-448.
- [104] S. Luca, D.V. Filippov, J.H. van Boom, H. Oschkinat, H.J. de Groot, and M. Baldus, Secondary chemical shifts in immobilized peptides and proteins: a qualitative basis for structure refinement under magic angle spinning. *J. Biomol. NMR* 20 (2001) 325-331.

- [105] H. Saitô, Conformation-dependent ^{13}C chemical shifts: A new means of conformational characterization as obtained by high-resolution solid-state ^{13}C NMR. *Magn. Reson. Chem.* 24 (1986) 835-852.
- [106] D.S. Wishart, B.D. Sykes, and N.J.O. Thomas L. James, Chemical shifts as a tool for structure determination, *Methods. Enzymol.*, Academic Press, 1994, pp. 363-392.
- [107] O.M. El-Agnaf, R. Jakes, M.D. Curran, D. Middleton, R. Ingenito, E. Bianchi, A. Pessi, D. Neill, and A. Wallace, Aggregates from mutant and wild-type alpha-synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of beta-sheet and amyloid-like filaments. *FEBS. Lett.* 440 (1998) 71-75.
- [108] O.M. El-Agnaf, and G.B. Irvine, Aggregation and neurotoxicity of alpha-synuclein and related peptides. *Biochem. Soc. Trans.* 30 (2002) 559-565.
- [109] B.I. Giasson, I.V. Murray, J.Q. Trojanowski, and V.M. Lee, A hydrophobic stretch of 12 amino acid residues in the middle of alpha-synuclein is essential for filament assembly. *J. Biol. Chem.* 276 (2001) 2380-2386.
- [110] F. Shewmaker, D. Kryndushkin, B. Chen, R. Tycko, and R.B. Wickner, Two prion variants of Sup35p have in-register parallel beta-sheet structures, independent of hydration. *Biochemistry* 48 (2009) 5074-5082.
- [111] M.J. Bayro, G.T. Debelouchina, M.T. Eddy, N.R. Birkett, C.E. MacPhee, M. Rosay, W.E. Maas, C.M. Dobson, and R.G. Griffin, Intermolecular structure determination of amyloid fibrils with magic-angle spinning and dynamic nuclear polarization NMR. *J. Am. Chem. Soc.* 133 (2011) 13967-13974.
- [112] G.T. Debelouchina, G.W. Platt, M.J. Bayro, S.E. Radford, and R.G. Griffin, Intermolecular Alignment in beta(2)-Microglobulin Amyloid Fibrils. *J. Am. Chem. Soc.* 132 (2010) 17077-17079.
- [113] J.J. Helmus, K. Surewicz, M.I. Apostol, W.K. Surewicz, and C.P. Jaronec, Intermolecular alignment in Y145Stop human prion protein amyloid fibrils probed by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 133 (2011) 13934-13937.
- [114] J. Yang, M.L. Tasayco, and T. Polenova, Magic angle spinning NMR experiments for structural studies of differentially enriched protein interfaces and protein assemblies. *J. Am. Chem. Soc.* 130 (2008) 5798-5807.
- [115] A.J. Nieuwkoop, and C.M. Rienstra, Supramolecular protein structure determination by site-specific long-range intermolecular solid state NMR spectroscopy. *J. Am. Chem. Soc.* 132 (2010) 7570-7571.
- [116] R.B. Wickner, F. Dyda, and R. Tycko, Amyloid of Rnq1p, the basis of the [PIN+] prion, has a parallel in-register beta-sheet structure. *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 2403-2408.
- [117] P. Güntert, W. Braun, M. Billeter, and K. Wüthrich, Automated Stereospecific ^1H NMR Assignments and Their Impact on the Precision of Protein Structure Determinations in Solution. *J. Am. Chem. Soc.* 111 (1989) 3997-4004.
- [118] D. Neri, T. Szyperski, G. Otting, H. Senn, and K. Wüthrich, Stereospecific Nuclear Magnetic Resonance Assignments of the Methyl Groups of Valine and

References

- Leucine in the DNA-Binding Domain of the 434 Repressor by Biosynthetically Directed Fractional ^{13}C Labeling. *Biochemistry* 28 (1989) 7510-7516.
- [119] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono, and P. Guntert, Optimal isotope labelling for NMR protein structure determinations. *Nature* 440 (2006) 52-57.
- [120] P.C. Driscoll, A.M. Gronenborn, and G.M. Clore, The influence of stereospecific assignments on the determination of three-dimensional structures of proteins by nuclear magnetic resonance spectroscopy: Application to the sea anemone protein BDS-I. *FEBS. Lett.* 243 (1989) 223-233.
- [121] J. Janin, S. Miller, and C. Chothia, Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* 204 (1988) 155-164.
- [122] D.F. Hansen, and L.E. Kay, Determining Valine Side-Chain Rotamer Conformations in Proteins from Methyl ^{13}C Chemical Shifts: Application to the 360 kDa Half-Proteasome. *J. Am. Chem. Soc.* 133 (2011) 8272-8281.
- [123] M. Hong, T.V. Mishanina, and S.D. Cady, Accurate Measurement of Methyl ^{13}C Chemical Shifts by Solid-State NMR for the Determination of Protein Side Chain Conformation: The Influenza A M2 Transmembrane Peptide as an Example. *J. Am. Chem. Soc.* 131 (2009) 7806-7816.
- [124] A. Loquet, N.G. Sgourakis, R. Gupta, K. Giller, D. Riedel, C. Goosmann, C. Griesinger, M. Kolbe, D. Baker, S. Becker, and A. Lange, Atomic model of the type III secretion system needle. *Nature* 486 (2012) 276-279.
- [125] H.S. Atreya, and K.V. Chary, Selective 'unlabeling' of amino acids in fractionally ^{13}C labeled proteins: An approach for stereospecific NMR assignments of CH_3 groups in Val and Leu residues. *J. Biomol. NMR.* 19 (2001) 267-272.
- [126] G. Ostler, A. Soteriou, C.M. Moody, J.A. Khan, B. Birdsall, M.D. Carr, D.W. Young, and J. Feeney, Stereospecific assignments of the leucine methyl resonances in the ^1H NMR spectrum of *Lactobacillus casei* dihydrofolate reductase. *FEBS. Lett.* 318 (1993) 177-180.
- [127] M.J. Plevin, O. Hamelin, J. Boisbouvier, and P. Gans, A simple biosynthetic method for stereospecific resonance assignment of prochiral methyl groups in proteins. *J. Biomol. NMR.* 49 (2011) 61-67.
- [128] V. Tugarinov, and L.E. Kay, Stereospecific NMR Assignments of Prochiral Methyls, Rotameric States and Dynamics of Valine Residues in Malate Synthase G. *J. Am. Chem. Soc.* 126 (2004) 9827-9836.
- [129] P. Gans, O. Hamelin, R. Sounier, I. Ayala, M.A. Durá, C.D. Amero, M. Noireclerc-Savoye, B. Franzetti, M.J. Plevin, and J. Boisbouvier, Stereospecific Isotopic Labeling of Methyl Groups for NMR Spectroscopic Studies of High-Molecular-Weight Proteins. *Angew. Chem. Int. Ed. Engl.* 49 (2010) 1958-1962.
- [130] H. Senn, B. Werner, B.A. Messerle, C. Weber, R. Traber, and K. Wüthrich, Stereospecific assignment of the methyl ^1H NMR lines of valine and leucine in polypeptides by nonrandom ^{13}C labeling. *FEBS. Lett.* 249 (1989) 113-118.
- [131] D. Neri, G. Otting, and K. Wüthrich, ^1H and ^{13}C NMR Chemical Shifts of the Diastereotopic Methyl Groups of Valyl and Leucyl Residues in Peptides and Proteins. *Tetrahedron* 46 (1990) 3287-3296.

- [132] M. Schubert, T. Manolikas, M. Rogowski, and B.H. Meier, Solid-state NMR spectroscopy of 10% ^{13}C labeled ubiquitin: spectral simplification and stereospecific assignment of isopropyl groups. *J. Biomol. NMR.* 35 (2006) 167-173.
- [133] D.H.G. Crout, C.J.R. Hedgecock, E.L. Lipscomb, and F.B. Armstrong, Stereochemistry of Valine Biosynthesis. *Eur. J. Biochem.* 110 (1980) 439-444.
- [134] R.K. Hill, and S.-J. Yan, Stereochemistry of Valine and Isoleucine Biosynthesis II. Absolute Configuration of (-) α,β -Dihydroxyisovaleric Acid and (-) α,β -Dihydroxy- β -Methylvaleric Acid. *Bioorg. Chem.* 1 (1971) 446-456.
- [135] R.K. Hill, S. Yan, and S.M. Arfin, Stereochemistry of valine and isoleucine biosynthesis. III. Enzymic discrimination between diastereotopic enol faces in the dehydrase step of valine biosynthesis. *J. Am. Chem. Soc.* 95 (1973) 7857-7859.
- [136] R.K. Hill, S. Sawada, and S.M. Arfin, Stereochemistry of Valine and Isoleucine Biosynthesis: IV. Synthesis, Configuration, and Enzymatic specificity of α -Acetolactate and α -Aceto- α -hydroxybutyrate. *Bioorg. Chem.* 8 (1979) 175-189.
- [137] S.R. Sylvester, and C.M. Stevens, Stereospecificity of the Reductoisomerase-Catalyzed Step in the Pathway of Biosynthesis of Valine and Leucine. *Biochemistry* 18 (1979) 4529-4531.
- [138] G. Lv, A. Kumar, K. Giller, M.L. Orcellet, D. Riedel, C.O. Fernandez, S. Becker, and A. Lange, Structural Comparison of Mouse and Human α -synuclein Amyloid Fibrils by Solid-State NMR. *J. Mol. Biol.* 420 (2012) 99-111.
- [139] K. Seidel, M. Etzkorn, H. Heise, S. Becker, and M. Baldus, High-Resolution Solid-State NMR Studies on Uniformly [$^{13}\text{C},^{15}\text{N}$]-Labeled Ubiquitin. *Chembiochem* 6 (2005) 1638-1647.
- [140] A. Böckmann, C. Gardiennet, R. Verel, A. Hunkeler, A. Loquet, G. Pintacuda, L. Emsley, B. Meier, and A. Lesage, Characterization of different water pools in solid-state NMR protein samples. *J. Biomol. NMR.* 45 (2009) 319-327.
- [141] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6 (1995) 277-293.
- [142] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, and E.D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59 (2005) 687-696.
- [143] O. Poyraz, H. Schmidt, K. Seidel, F. Delissen, C. Ader, H. Tenenboim, C. Goosmann, B. Laube, A.F. Thunemann, A. Zychlinsky, M. Baldus, A. Lange, C. Griesinger, and M. Kolbe, Protein refolding is required for assembly of the type three secretion needle. *Nat. Struct. Mol. Biol.* 17 (2010) 788-792.
- [144] E.M.F. Brekke, A.B. Walls, A. Schousboe, H.S. Waagepetersen, and U. Sonnewald, Quantitative importance of the pentose phosphate pathway determined by incorporation of ^{13}C from [2- ^{13}C]- and [3- ^{13}C]glucose into TCA cycle intermediates and neurotransmitter amino acids in functionally intact neurons. *J. Cereb. Blood. Flow. Metab.* 32 (2012) 1788-1799.
- [145] P. Lundström, D.F. Hansen, and L.E. Kay, Measurement of carbonyl chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy:

References

- comparison between uniformly and selectively ^{13}C labeled samples. *J. Biomol. NMR.* 42 (2008) 35-47.
- [146] C. Guo, and V. Tugarinov, Selective ^1H - ^{13}C NMR spectroscopy of methyl groups in residually protonated samples of large proteins. *J. Biomol. NMR.* 46 (2010) 127-133.
- [147] T. Szyperski, J.E. Bailey, and K. Wuhrich, Detecting and dissecting metabolic fluxes using biosynthetic fractional ^{13}C labeling and two-dimensional NMR spectroscopy. *Trends. Biotechnol.* 14 (1996) 453-459.

Acknowledgements

First, I would like to thank my supervisor Dr. Adam Lange for offering me this invaluable opportunity to work at the Max Planck Institute for Biophysical Chemistry (MPIBPC) on the interesting projects. Thanks a lot for your supervision, illuminating discussions, and all continuous supports.

I also thank Prof. Dr. Christian Griesinger at the MPIBPC for providing excellent facilities, outstanding research environment in the department for NMR-based Structural Biology.

I would like to thank Prof. Dr. Kai Tittmann from the Georg-August-Universität Göttingen and Prof. Berend de Groot from the MPIBPC for their membership of my thesis committee and helpful suggestions during the committee meetings.

I would like to thank all my current and former colleagues for their assistance and cooperation, for helping me learn solid-state NMR and for a friendly working atmosphere. In particular, I thank Dr. Ashutosh Kumar for the instructions of solid-state NMR experiments, for introducing the project of the structural characterization of mouse α -synuclein fibrils, and for the helpful discussion regarding the sequential resonance assignment at the beginning. I thank Dr. Antoine Loquet for the collaboration on the projects of sparse labeling schemes and the stereospecific assignment. I also thank Hannes Klaus Fasshuber for the cooperation on the stereospecific assignment project. I thank Jean-Philippe Demers and Dr. Zrinka Gattin for the academic discussion and kind help in many aspects of my work and life. I also thank Gitta Angerstein for her help and support in the lab and in all matters of daily life and work at the institute. I thank Dr. Veniamin Shevelkov, Dr. Ewewei Saw, Dr. Guowei Yin, Harindranath Kadavath for their help in my thesis writing.

I also thank our secretaries Mrs. Breiner and Mrs. Silberer, our IT system administrators Dr. Dirk Bockelmann and Jürgen Arve for their daily supports.

Acknowledgements

I would like to thank Dr. Stefan Becker and Karin Giller at the MPIBPC for their expert sample preparations of mouse α -synuclein, as well as Dr. Dietmar Riedel and Gudrun Heim at the MPIBPC for performing negative stained electron microscopy mouse α -synuclein samples.

I also thank Christina Bach, Kirsten Pöhlker and other persons in the GGNB team for their kind help and excellent organization of the course and academic activities.

I thank all my dear friends in Goettingen, especially the friends at the MPIBPC for your companionship and friendship.

Last, but not least, I would like to give my infinite thanks to my parents for their supports and understandings, and Yun Huang for his love and companionship.

Curriculum Vitae

Personal data

Name Guohua Lv
Gender Female
Date and place of birth Mar 15, 1984, Shandong, China
Email glv@nmr.mpibpc.mpg.de
Current Address Fassberg 11, 37077, Goettingen, Germany
Tel +49 17699926768

Education and Position

08/2009 – **Max Planck Institute for Biophysical Chemistry**,
Goettingen Germany
PhD student in the department of NMR-based Structural
Biology
Supervisor: Dr. Adam Lange
Dissertation: Protein Structure Characterization by Solid-State
NMR: Structural Comparison of Mouse and Human α -
Synuclein Fibrils, Sparse ^{13}C Labeling Schemes, and
Stereospecific Assignment of Val and Leu Prochiral Methyl
Groups

09/2006 – 06/2009 **Lanzhou University**, Lanzhou, China
M.S. in Biochemistry and Molecular Biology
Supervisor: Prof. Shouliang Dong
Thesis: Solution Structural Determination and Dynamics
Investigation of BAM8-22 and truncated peptides of BAM8-
22 by NMR Spectroscopy and Molecular Dynamics
Simulation

09/2002 – 06/2006 **Lanzhou University**, Lanzhou, China
B.S. Biology Science

Publications

Guohua Lv, Hannes Klaus Faßhuber, Antoine Loquet, Jean-Philippe Demers, Vinesh Vijayan, Karin Giller, Stefan Becker, and Adam Lange* A Straightforward Method for Stereospecific Assignment of Val and Leu Prochiral Methyl Groups by Solid-State NMR: Scrambling in the [2-¹³C]Glucose Labeling Scheme. *J. Magn. Reson.* 2013 Mar 29;228:45-49. (Cover Article)

Guohua Lv, Ashutosh Kumar, Karin Giller, Maria L. Orcellet, Dietmar Riedel, Claudio O. Fernández, Stefan Becker, and Adam Lange*. Structural Comparison of Mouse and Human α -Synuclein Amyloid Fibrils by Solid-State NMR. *J. Mol. Biol.* 2012 Jun 29;420(1-2):99-111.

Antoine Loquet, **Guohua Lv**, Karin Giller, Stefan Becker, and Adam Lange*. ¹³C Spin Dilution for Simplified and Complete Solid-State NMR Resonance Assignment of Insoluble Biological Assemblies. *J. Am. Chem. Soc.* 2011 Apr 6;133(13):4722-4725.

Guohua Lv, Shouliang Dong*. Solution Structural Investigation and Conformation-Activity Relationship of BAM8-22 by NMR and Molecular Dynamics Simulations. *Bioorg. Med. Chem. Lett.* 2010 Feb 1;20(3):1260-1262.