

# **Magic-angle spinning NMR investigation of the interaction between FAT10 and NUB1L**

**Doctoral thesis for obtaining the  
academic degree**

**Doctor of Natural Sciences (Dr. rer. nat.)**

submitted by

Weiss, Charlotte

at the

Universität  
Konstanz



Faculty of Sciences

Department of Chemistry

Konstanz, 2026

*Konstanzer Online-Publikations-System (KOPS)*  
URL: <http://nbn-resolving.de/urn:nbn:de:bsz:352-2-zr19ac0ym11c3>

Date of the oral examination: 13 November 2025

1. Reviewer: Dr. Guinevere Mathies

2. Reviewer: Dr. Annette Aichem

## Abstract

Proteasomal degradation is generally associated with the ubiquitin-proteasome-system. However, numerous ubiquitin-independent proteasomal degradation pathways exist with underestimated relevance for proteostasis. Under inflammatory conditions, FAT10 expression is upregulated targeting substrate proteins for rapid and irreversible degradation by the 26S proteasome. FAT10 is a member of the ubiquitin-like modifier family and consists of two domains (N- and C-domain). In contrast to the heat-stable ubiquitin, the biological function of FAT10 is determined by its weak folding. To understand how the weak folding of FAT10 enables rapid degradation and which role the interaction with its binding partner NUB1L is playing, structural insights at atomic scale are indispensable.

Chapters 2 and 3 of this thesis show the strength of MAS NMR in the investigation of FAT10. For the isolated, cysteine-free N-domain of FAT10 (N-FAT10-C0), well-resolved spectra and a nearly complete site-specific assignment of carbon and nitrogen resonances are obtained confirming the feasibility of a MAS NMR study. Spectra of N-FAT10-C0 in complex with NUB1L are of good quality, but differ substantially from the isolated N-domain. Sequential assignments are limited to a stretch of residues that interact with NUB1L in the form of an intermolecular  $\beta$ -sheet, while the rest of the N-domain of FAT10 is unstructured. The observation of the N-terminal residue of FAT10 assumes a docking to NUB1L which presumably presents the largely unfolded N-domain to the proteasome for rapid degradation. In case of the isolated, wild-type N-domain of FAT10 (N-FAT10-WT), sample preparation is complicated as the protein starts to aggregate upon concentration. The structural heterogeneity is reflected in broad spectral features. Co-sedimentation with NUB1L changes the scene dramatically: The increased resolution allows for the characterization of the interaction between N-FAT10-WT and NUB1L, which is identical to N-FAT10-C0 in complex with NUB1L. Thus, co-sedimentation with known binding partners is considered as a general strategy for the sample preparation of weakly folded proteins.

## Zusammenfassung

Der Abbau von Proteinen durch das Proteasom wird allgemein mit dem Ubiquitin-Proteasom-System in Verbindung gebracht. Es existieren jedoch zahlreiche Ubiquitin-unabhängige Abbauewege, deren Bedeutung für die Proteostase unterschätzt wird. Unter entzündlichen Bedingungen wird die Expression von FAT10 hochreguliert, wodurch Substratproteine für einen schnellen und irreversiblen Abbau durch das 26S Proteasom bestimmt sind. FAT10 gehört zur Familie der Ubiquitin-ähnlichen Modifikatoren und besteht aus zwei Domänen (N- und C-Domäne). Im Gegensatz zum hitzestabilen Ubiquitin ist die biologische Funktion von FAT10 durch seine lose Faltung bestimmt. Um zu verstehen, wie die lose Faltung von FAT10 einen schnellen Abbau ermöglicht und welche Rolle die Interaktion mit seinem Bindungspartner NUB1L spielt, sind strukturelle Einblicke auf atomarer Ebene unerlässlich.

Die Kapitel 2 und 3 dieser Arbeit zeigen die Stärke der MAS-NMR-Spektroskopie bei der Untersuchung von FAT10. Für die isolierte, Cystein-freie N-Domäne von FAT10 (N-FAT10-C0) konnten gut aufgelöste Spektren und eine nahezu vollständige, sequenzspezifische Zuordnung von Kohlenstoff- und Stickstoffresonanzen erhalten werden, was die Durchführbarkeit einer MAS-NMR-Studie bestätigt. Die Spektren von N-FAT10-C0 im Komplex mit NUB1L sind von guter Qualität, unterscheiden sich jedoch erheblich von der isolierten N-Domäne. Sequentielle Zuordnungen beschränken sich auf einen Abschnitt von Aminosäuren, die mit NUB1L in Form eines intermolekularen  $\beta$ -Faltblatts interagieren, während der restliche Teil der N-Domäne von FAT10 unstrukturiert ist. Die Beobachtung des N-Terminus von FAT10 lässt auf ein Andocken an NUB1L schließen. Es kann angenommen werden, dass NUB1L die weitgehend entfaltete N-Domäne dem Proteasom präsentiert, was zum schnellen Abbau führt. Im Fall der isolierten Wildtyp N-Domäne von FAT10 (N-FAT10-WT) ist die Probenpräparation erschwert, da das Protein bei der Konzentrierung zu aggregieren beginnt. Die strukturelle Heterogenität spiegelt sich in breiten Linien wider. Die gemeinsame Sedimentation mit NUB1L verändert das Bild dramatisch: Die erhöhte Auflösung ermöglicht die Charakterisierung der Interaktion zwischen N-FAT10-WT und NUB1L, die identisch mit N-FAT10-C0 im Komplex mit NUB1L ist. Daher gilt die Sedimentation mit bekannten Bindungspartnern als allgemeine Strategie für die Probenpräparation schwach gefalteter Proteine.

---

# Table of Contents

<b>Abstract</b> .....	ii
<b>Zusammenfassung</b> .....	iii
<b>Table of Contents</b> .....	iv
<b>1 Chapter 1 Introduction</b> .....	1
<b>1.1 Why use MAS NMR for biosolids?</b> .....	1
<b>1.2 History of solid-state NMR</b> .....	2
1.2.1 Anisotropic interactions and magic-angle spinning .....	2
1.2.2 Cross-polarization of low- $\gamma$ nuclei.....	4
1.2.3 Heteronuclear decoupling .....	4
1.2.4 Recoupling and distance measurements .....	5
1.2.5 Resonance assignment methods.....	6
<b>1.3 Proteasomal degradation pathways in eukaryotic cells</b> .....	8
<b>1.4 Thesis outline</b> .....	11
<b>2 Chapter 2 Intermolecular <math>\beta</math>-sheet formation guides interaction between FAT10 and NUB1L</b> .....	12
<b>2.1 Abstract</b> .....	12
<b>2.2 Introduction</b> .....	12
<b>2.3 Materials and methods</b> .....	13
<b>2.4 Results</b> .....	16
2.4.1 Valuation of Samples .....	16
2.4.2 Resonance Assignments and Secondary Structure of Microcrystalline N-FAT10-C0.....	17
2.4.3 Resonance Assignments and Secondary Structure of Lyophilized/Rehydrated N-FAT10-C0.....	18
2.4.4 Interaction of N-FAT10-C0 and NUB1L.....	21
<b>2.5 Discussion</b> .....	23
<b>2.6 Conclusion</b> .....	25
<b>2.7 Supporting Information</b> .....	25
2.7.1 Protein expression and purification .....	25
2.7.2 Preparation of samples for MAS NMR.....	27
2.7.3 MAS NMR spectroscopy.....	28
2.7.4 $^1\text{H}$ - $^{15}\text{N}$ cross-polarization spectra.....	31
2.7.5 $^{13}\text{C}$ - $^{13}\text{C}$ and $^{15}\text{N}$ - $^{13}\text{C}$ spectra of microcrystalline N-FAT10-C0 and visualization of secondary structure predictions.....	32
2.7.6 Strip plot of microcrystalline N-FAT10-C0.....	34

---

2.7.7	Strip plots of lyophilized/rehydrated N-FAT10-C0 .....	35
2.7.8	Torsion angle analysis of lyophilized/rehydrated N-FAT10-C0 .....	37
2.7.9	Temperature dependence of N-FAT10-C0 in complex with NUB1L ...	38
2.7.10	Strip plot of N-FAT10-C0 in complex with NUB1L .....	39
2.7.11	Torsion angle analysis of N-FAT10-C0 in complex with NUB1L .....	40
2.7.12	AlphaFold-Multimer prediction for the interaction of N-FAT10-C0 and NUB1L.....	41
<b>3</b>	<b>Chapter 3 Co-sedimentation as the key to the investigation of wild-type FAT10 by MAS NMR spectroscopy .....</b>	<b>42</b>
<b>3.1</b>	<b>Abstract .....</b>	<b>42</b>
<b>3.2</b>	<b>Introduction .....</b>	<b>42</b>
<b>3.3</b>	<b>Materials and methods.....</b>	<b>44</b>
<b>3.4</b>	<b>Results .....</b>	<b>46</b>
<b>3.5</b>	<b>Discussion .....</b>	<b>51</b>
<b>3.6</b>	<b>Conclusion.....</b>	<b>53</b>
<b>3.7</b>	<b>Supporting Information .....</b>	<b>54</b>
3.7.1	Protein expression and purification .....	54
3.7.2	Preparation of samples for MAS NMR.....	56
3.7.3	MAS NMR spectroscopy.....	56
3.7.4	$^1\text{H}$ - $^{15}\text{N}$ cross-polarization spectra.....	59
3.7.5	Strip plot of N-FAT10-WT in complex with NUB1L .....	60
3.7.6	$^{15}\text{N}$ - $^{13}\text{C}$ spectra of N-FAT10-WT in complex with NUB1L .....	61
3.7.7	Torsion angle analysis of N-FAT10-WT in complex with NUB1L .....	62
	<b>References .....</b>	<b>63</b>
	<b>Curriculum Vitae.....</b>	<b>71</b>

# 1 Chapter 1 Introduction

## 1.1 Why use MAS NMR for biosolids?

Biological macromolecules form the basis of cellular function. One macromolecular group is constituted by proteins: They provide structural support, work in ion transport and cellular signalling and are functional in metabolic pathways and reactions of the immune system. Proteins are built of 20 canonical amino acids, which are organized in four structural levels.<sup>1</sup> The primary structure refers to the linear sequence of amino acids. The regular local structure of the backbone characterized by defined regions in the Ramachandran plot<sup>2</sup> (the graphical representation of  $\Phi$  and  $\Psi$  backbone torsion angles) is known as secondary structure including  $\alpha$ -helices, extended  $\beta$ -strands forming parallel or anti-parallel  $\beta$ -sheets and  $\beta$ -turns. The 3-dimensional folding or tertiary structure of a protein corresponds to the arrangement of secondary structure elements in space and is determined by interactions of the side chains (hydrophobic interactions, van der Waals interactions, hydrogen bonds, salt bridges and disulfide bonds). The association of two or multiple polypeptide chains in a single functional protein is described in the quaternary structure.<sup>1</sup>

Similar to the different structural levels, the protein's sequence, structure, dynamics and function are organized hierarchically. Unless a sequence homolog with functional annotation is available, protein function is hardly predicted from the sequence alone. As whole genomes can be sequenced readily nowadays,<sup>3</sup> the gap between the tremendous number of novel genes and functional annotation is growing. Instead of the sequence, structural homologs predict more reliable the protein function. Dynamics ultimately relates structure and function<sup>4</sup> emphasizing the importance of techniques for the structural and dynamical characterization to bridge the gap between sequence and function space.

The first 3-dimensional structure of a protein was solved by X-ray crystallography<sup>5</sup> setting the scene to develop into the main technique for structural biology.<sup>6,7</sup> The majority of deposited protein structures is determined by X-ray crystallography and hence, protein structure prediction methods such as the artificial intelligence system AlphaFold<sup>8</sup> are mostly trained on experimental X-ray structures. By detecting the angles and intensities of X-rays diffracted by a good quality crystal (a so-called diffraction pattern), the electron density and thus the positions of the atoms within the molecule can be inferred. As protons, comprising one electron, are poor X-ray scatterers, structural information is mainly restricted to heavy atoms. Moreover, the static nature of crystals, where molecules are in fixed regular positions, together with the use of cryogenic temperatures to reduce radiation damage and increase the intensity of the diffracted beams complicate the investigation of dynamics.

An emerging technique in the field of structural biology is cryo-electron microscopy (cryo-EM).<sup>9</sup> The macromolecule or macromolecular assembly to be investigated is vitrified on the cryo-EM grid to prevent the formation of ice crystals preserving the native state of the sample. Datasets are recorded on a transmission electron microscope providing structural and dynamic information

complementary to X-ray crystallography. A wide range of samples of different sizes can be studied, but often resolution is limited impeding functional conclusions.

Using solution-state nuclear magnetic resonance (NMR) spectroscopy, highly soluble systems can be studied in near-physiological environments and temperatures. Structural elucidation, the investigation of binding events and the identification of molecular motion are possible at atomic resolution inclusive of protons. In fact,  $^1\text{H}$  represents the NMR-active hydrogen species and is the most common nucleus in biomolecules in almost 100 % natural abundance (NA).<sup>10</sup> Limitations of the method arise facing large molecular systems as rotational correlation times become longer, reduced solubility or assembly formation leading to aggregation or membrane proteins that naturally are found in lipid bilayers. Additionally, details about motions in the nanosecond to microsecond regime and the structure of the hydrogen bonding network are difficult to determine by solution-state NMR.<sup>10</sup>

By magic-angle spinning (MAS) NMR spectroscopy, boundaries of classical methods (X-ray crystallography and solution-state NMR) can be pushed. Structural, dynamic and hence functional information gets accessible for solid and semi-solid systems covering insoluble or aggregated molecules, micro- and nanocrystalline as well as amorphous samples, complexes co-sedimented from solution, higher-order assemblies like ribosomes, proteasomes or viruses, membrane proteins, biominerals and organelles, even in a cellular context.<sup>10</sup> Beside molecular topology and chemical structure, e.g. oxidization or protonation states, small distances important in ligand binding or protein-protein interactions can be measured.<sup>11</sup> In the solid-state, molecular tumbling is absent. This allows for the investigation of molecular systems independent of size and motions on all time scales at atomic resolution. At the same time, anisotropic interactions are prominent causing severe line broadening. In order to still get interpretable spectra like in solution, many developments in hardware, design, set-up and methodology have been accomplished in the history of solid-state NMR; one is magic-angle spinning.

## 1.2 History of solid-state NMR

### 1.2.1 Anisotropic interactions and magic-angle spinning

Nuclear magnetic resonance was first described by Rabi *et al.* in 1938.<sup>12</sup> Already in the earliest experiments, it became apparent that liquids generally show sharp resonances, whereas the ones of solids are much broader.<sup>13</sup> In solution, molecules can randomly rotate and translate. In this way, anisotropic, i.e. orientation-dependent interactions are averaged. Molecular tumbling is impeded in condensed matter and a more thorough analysis of anisotropic interactions is needed to understand the origin of broad lines in spectra of solids. The focus is set on spin  $I = 1/2$  nuclei like  $^1\text{H}$ ,  $^{13}\text{C}$  or  $^{15}\text{N}$  and therefore, the anisotropy of chemical shift and magnetic dipolar interactions exclusive of quadrupolar interactions involving spin  $I \geq 1$  nuclei are discussed subsequently.

The application of a strong, external magnetic field  $B_0$  not only lifts the degeneracy of the nuclear spin states (Zeeman interaction), but also induces circulating currents of the surrounding electrons. The induced currents oppose  $B_0$  and modulate the local magnetic field sensed by a given nucleus. This

shielding effect of the electrons makes NMR a sensitive method to probe the different chemical environments, in which nuclei of the same isotope can be found in. Depending on the orientation of the molecule with respect to  $B_0$ , the degree to which the chemical shielding affects the resonance frequency of a given nucleus varies. Consequently, in a solid powder containing randomly arranged crystallites, all possible molecular orientations are sampled giving rise to a spectrum with severe line broadening. Each nuclear spin possesses a magnetic moment which produces a small magnetic field. Due to the Zeemann interaction, the energy levels of the nuclear spins aligned either parallel (spin-up) or anti-parallel (spin-down) to  $B_0$  are split. Depending on the alignment to  $B_0$ , the small nuclear magnetic fields either add up or subtract from the local magnetic field of dipolar coupled, close-by nuclei ( $< 10 \text{ \AA}$ ). For two spins  $I$  and  $S$  (commonly,  $I$  refers to abundant spins like  $^1\text{H}$  and  $S$  to rare spins like  $^{13}\text{C}$  or  $^{15}\text{N}$ ) at the internuclear distance  $r_{IS}$ , the size of the dipolar coupling  $d_{IS}$  is given by the following equation<sup>14</sup>:

$$(1) \quad d_{IS} = -(\mu_0 \gamma_I \gamma_S \hbar / 8\pi r_{IS}^3)(3\cos^2\theta - 1)$$

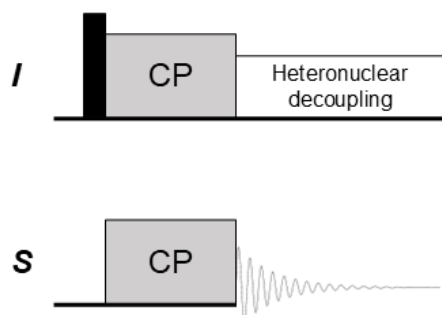
$\mu_0$  corresponds to the permeability of free space,  $\gamma_I \gamma_S$  to the product of the gyromagnetic ratios of  $I$  and  $S$  and  $\theta$  specifies the angle of the internuclear vector relative to  $B_0$ . Nuclei in close proximity show a strong coupling which falls off rapidly with the cube of the internuclear distance. As dipolar coupling is mediated through space instead of bonding electrons, nuclei within different molecules can interact with each other. The same applies to nuclei in proteins that are separated sequentially, but are brought together upon protein folding. The effect of dipolar coupling on the nuclear resonance frequencies is determined beside internuclear distance by the orientation-dependent term  $(3\cos^2\theta - 1)$ . The same term describes the orientation dependence of chemical shift anisotropy (CSA).

To imitate nature of molecular tumbling in solution, it is unfeasible to rotate a solid sample around several axes at high speed. Alternatively, Andrew *et al.* (1959) and Lowe (1959) could demonstrate theoretically and experimentally that upon rotation of solid specimens around an axis inclined at  $\theta = 54.74^\circ$  with respect to the static magnetic field  $B_0$ , line broadening due to anisotropic interactions can be removed.<sup>15,16</sup> This angle of  $\theta = 54.74^\circ$  is referred to as the magic-angle. The practice of magic-angle spinning results in a spectrum determined by isotropic chemical shifts and spinning sidebands at integer multiples of the rotation rate  $\omega_r$ .

Protons have the highest gyromagnetic ratio among nuclei and are highly abundant in biological macromolecules resulting in a strong  $^1\text{H}$ - $^1\text{H}$  dipolar coupling network. Incomplete averaging of the homonuclear dipolar coupling under MAS at moderate rates ( $< 40 \text{ kHz}$  that falls below the coupling between two proton spins reaching up to  $100 \text{ kHz}$ ) leads to broad proton lines. Thus, proton detection at moderate rates requires deuteration of the sample and subsequent back-exchange of a small number of labile protons to reduce proton density (so-called spin dilution).<sup>17</sup> Alternatively, low- $\gamma$  nuclei such as  $^{13}\text{C}$  or  $^{15}\text{N}$  are commonly detected in the solid-state using  $3.2 \text{ mm}$  E-free HCN probes capable of spinning up to  $24 \text{ kHz}$ . E-free (E refers to the electric field) HCN probes are ideally for measurements of biological relevant systems like proteins primarily composed of protons, carbons and nitrogens and aim to reduce radio frequency (RF) induced heating of wet samples.<sup>18,19</sup>

1.2.2 Cross-polarization of low- $\gamma$  nuclei

To compensate for the low sensitivity, the majority of solid-state NMR experiments begins with an initial polarization transfer from abundant proton to rare nuclei. Hartmann-Hahn cross-polarization (HHCP) or short CP is a well-known and routinely used technique for dipolar polarization transfer between heteronuclear spins introduced by Hartmann and Hahn in 1962.<sup>20</sup> The high polarization and short relaxation time of protons as well as the high resolution of rare nuclei such as  $^{13}\text{C}$  or  $^{15}\text{N}$



**Figure 1.1:** Pulse sequence for Hartmann-Hahn cross-polarization.  $I$  usually represents  $^1\text{H}$  and  $S$  low- $\gamma$  nuclei such as  $^{13}\text{C}$  or  $^{15}\text{N}$ . The filled rectangle defines a  $90^\circ$ -pulse. During the contact time, polarization is transferred from  $I$  to  $S$ . Heteronuclear decoupling is applied throughout the acquisition time.

are incorporated simultaneously in CP. The basic sequence is shown in Figure 1.1. At first, the  $I$  magnetization is brought into the  $xy$ -plane by a  $90^\circ$ -pulse (preparation pulse). During the CP contact time, a constant RF field is applied on the  $S$  spin while  $I$  is spin-locked. In this doubly rotating frame, polarization is transferred from the  $I$  to the  $S$  spin fulfilling the matching condition, i.e.  $|\omega_1^I| = |\omega_1^S| \Leftrightarrow |\gamma_I B_1^I| = |\gamma_S B_1^S|$ , where  $B_1$  corresponds to the magnetic field component of the respective RF fields.<sup>21</sup> Under MAS, the matching condition is modified to  $\omega_1^I \pm \omega_1^S = n\omega_r$  with  $n = 0, \pm 1, \pm 2, \dots$ . To account for fluctuations in experimental parameters, amplitude-modulated spin-lock pulses

including ramped or adiabatic-shaped RF fields are employed to guarantee for a stable CP condition.<sup>10</sup> Depending on the strength of the  $I$ - $S$  dipolar coupling, optimal contact times vary for every  $S$  spin and typically last from 100  $\mu\text{s}$  to 10 ms. For example, as backbone carbonyl carbons in proteins are not covalently bound to protons longer contact times than for aliphatic carbons are usually needed to build up polarization. By combining MAS with CP as well as heteronuclear decoupling, carbon (or nitrogen) spectra with high resolution and sensitivity can be obtained.

## 1.2.3 Heteronuclear decoupling

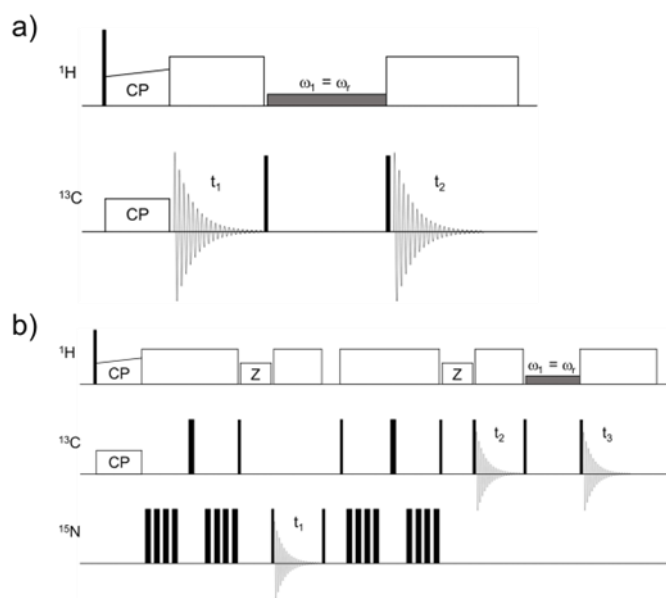
Homonuclear couplings of low- $\gamma$  nuclei are readily averaged under MAS. However, strong heteronuclear couplings involving proton nuclear spins such as  $^1\text{H}$ - $^{13}\text{C}$  or  $^1\text{H}$ - $^{15}\text{N}$  are still a nuisance even under MAS and require additional spin decoupling to reduce dephasing and hence line widths of the observed nuclei. Standardly, continuous-wave (CW) spin decoupling was used. The dipolar modulation of the resonance frequency of the  $S$  spin is opposite, depending on whether the  $I$  spin is aligned parallel (spin-up) or anti-parallel (spin-down) to the external magnetic field  $B_0$ . This means that upon CW irradiation, implying the application of a strong proton RF field, the  $I$  spin-state continuously changes between spin-up and spin-down eliminating heteronuclear coupling. Although heteronuclear decoupling by CW irradiation can be simply implemented without the needs for any optimization, it is not the most efficient way. Still, for a long time no alternative approach was available until in 1995, Bennett *et al.* introduced the first multiple-pulse technique, so-called two pulse phase modulation (TPPM), which outperformed CW decoupling.<sup>22</sup>

The TPPM scheme consists of pulses of length  $\tau_p$  with phases alternating between the values of  $-\frac{\phi}{2}$  and  $\frac{\phi}{2}$ .<sup>22</sup> Both, pulse lengths (close to the duration of a  $180^\circ$ -pulse) and phases (in the range of  $10$ - $25^\circ$ ) have to be optimized to accomplish good decoupling. The high performance of TPPM is limited to low MAS rates. At MAS frequencies above  $20$  kHz, TPPM is superseded by the X-inverse X (XiX) decoupling scheme<sup>23</sup>. The phase-inverted pulses in XiX are of  $1.85\tau_r$  or  $2.85\tau_r$  duration with  $\tau_r$  corresponding to the duration of a  $180^\circ$ -pulse. TPPM and XiX are just two of a large number of heteronuclear decoupling sequences developed over the last years including variants such as Frequency-Swept TPPM (SW<sub>r</sub>-TPPM)<sup>24</sup> or Small-Phase Incremental Alternation with 64 steps (SPINAL-64)<sup>25</sup>. On the one hand, anisotropic interactions are desired to average out through MAS and high-power heteronuclear decoupling. On the other hand, dipolar couplings and CSA contain useful structural information (internuclear distances, geometry of bonding) that can be reintroduced by different methods.

### 1.2.4 Recoupling and distance measurements

Selective recoupling is required to gain structural information based on the distance dependence of dipolar coupling. It is impractical to switch spinning on and off, so dipolar coupling under MAS is reintroduced by different means including couplings to abundant spins, the application of RF pulses or matching of the MAS spinning frequency to the difference in chemical shifts.<sup>26</sup>

The method reported by Gullion and Schaefer (1989) to measure the (weak) dipolar coupling between two nuclear species such as  $^{13}\text{C}$ - $^{15}\text{N}$  is called rotational echo double resonance (REDOR) and is widely used in structural studies of solids.<sup>27,28</sup> In an uniformly  $^{13}\text{C}$ , $^{15}\text{N}$ -labelled protein



**Figure 1.2:** Pulse sequences for  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  homo- and heteronuclear correlation spectroscopy. Narrow and wide filled rectangles correspond to  $\pi/2$ - and  $\pi$ -pulses, respectively. Open rectangles represent decoupling. (a) PDS and DARR. For DARR, protons are irradiated at an RF field of  $n$  times the spin rate  $\omega_r$  during mixing (here:  $\omega_1 = \omega_r$ ). (b) ZF TEDOR-DARR with  $\Delta$  as z filter period.

sample, the situation is more complex than for an isolated spin pair and a frequency selective version of REDOR (fsREDOR) is required to recouple a single  $^{13}\text{C}$ - $^{15}\text{N}$  dipolar interaction in a multiple spin system.<sup>29</sup> Alternatively, experiments that build on the closely related transferred-echo double-resonance (TEDOR) technique<sup>30,31</sup> allow for the simultaneous measurement of multiple  $^{13}\text{C}$ - $^{15}\text{N}$  distances in uniformly labelled samples. The approach of ZF TEDOR uses z-filter (ZF) periods to avoid the undesirable effects of  $^{13}\text{C}$ - $^{13}\text{C}$   $J$ -couplings that would otherwise disturb the measurement of weak  $^{13}\text{C}$ - $^{15}\text{N}$  dipolar couplings (corresponding to distances in the range of  $\sim 5$ - $6$  Å).<sup>32</sup> (The

pulse sequence of ZF TEDOR is shown in Figure 1.2b as part of the pulse sequence ZF TEDOR-DARR.)

Experiments that involve dipolar recoupling of homonuclear spins are particularly useful for the measurement of  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra as carbons are more sensitive than nitrogens and better resolved than protons at moderate MAS rates. Common pulse sequences with low power requirements enabling mixing times up to several hundreds of milliseconds are proton driven spin diffusion (PDS) and dipolar-assisted rotational resonance (DARR)<sup>34,35</sup> (Figure 1.2a). An extension of the concept of REDOR to the homonuclear case is realized e.g. in radio frequency driven dipolar recoupling (RFDR)<sup>36</sup>. Due to the attenuation of recoupled dipolar interactions between distant spins by the presence of stronger couplings to close-by spins (known as dipolar truncation)<sup>37</sup> and an assistance of protons depending on the recoupling mechanism of a certain pulse sequence, it is difficult to extract exact distances between coupled spins from carbon correlation spectra. Still, the first 3-dimensional structure of a protein in the solid-state was determined using PDS.<sup>38</sup>

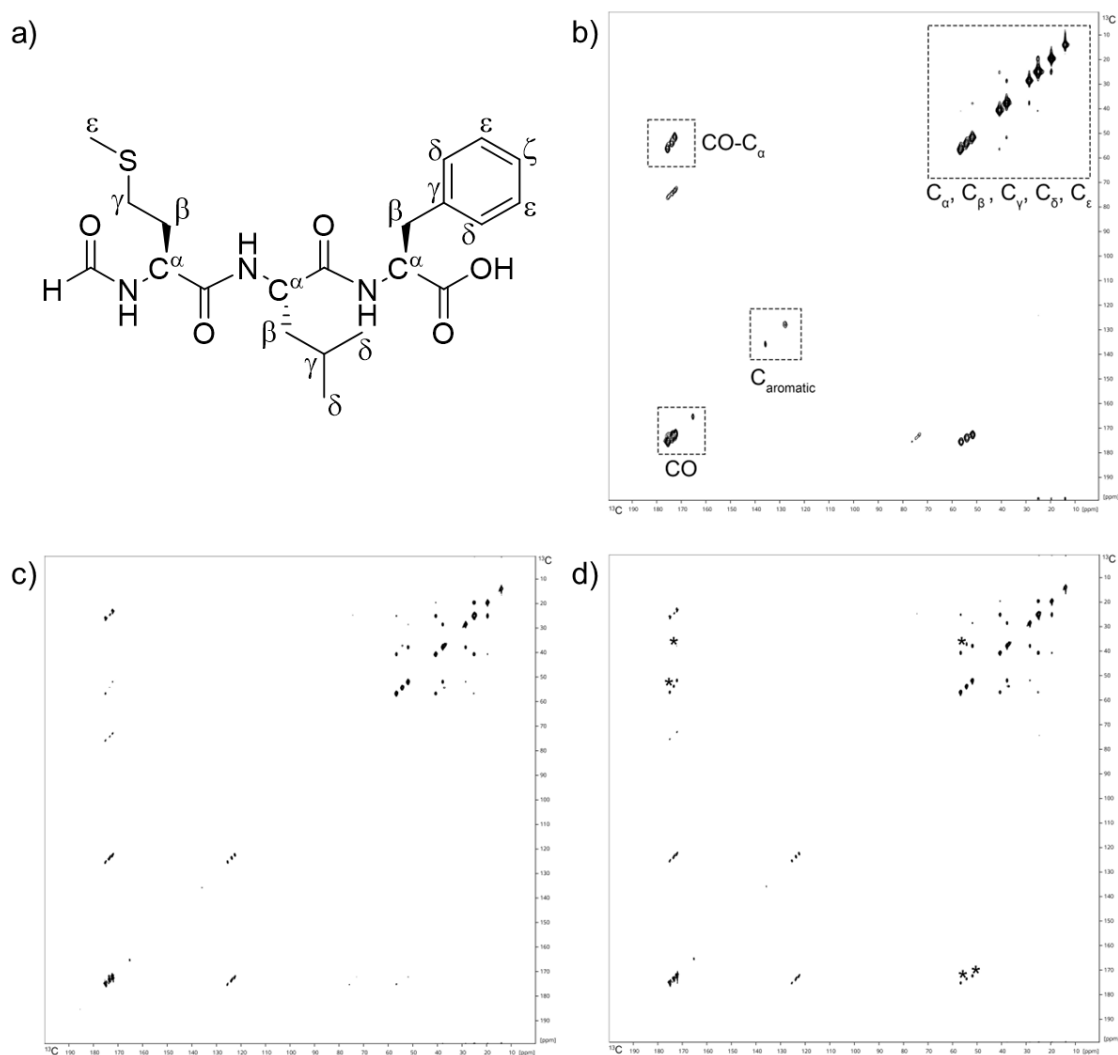
### 1.2.5 Resonance assignment methods

One core requirement for any structural, dynamic and functional study is the assignment of resonances to track processes site-specifically. In the liquid state, a standard protocol based on proton detection is established that allows for easy implementation and automation for the majority of  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled (protein) samples in solution.<sup>17</sup> In the solid-state, difficulties are encountered in proton detection, i.e. broad lines associated with strong proton dipolar couplings. Advances in the past ten years have made spinning speeds up to 100 kHz and most recently 160 kHz possible enabling the acquisition of resolved proton-detected spectra.<sup>39,40</sup>

At prevalent spinning speeds of 8-24 kHz, carbon-detected spectra are routinely recorded. The assignment of resonances *via* carbon-detection set to work on the identification of spin systems and corresponding amino acid types. For this purpose, 2D  $^{13}\text{C}$ - $^{13}\text{C}$ -correlation spectra (e.g. DARR, PDS or RFDR) are acquired and analysed providing information about sample quality, signal dispersion and spectral crowding at the same time.<sup>17</sup> The approach will be exemplified for the test sample of N-f-MLF-OH (10 % of uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled (excluding formyl carbonyls) N-f-MLF-OH diluted within natural abundance (NA) N-f-MLF-OH recrystallized from 2-propanol). The N-formylated tripeptide comprises the three amino acids methionine, leucine and phenylalanine (see Figure 1.3a) and is involved in chemotaxis.<sup>41,42</sup>

At low mixing times mainly one-bond  $^{13}\text{C}$ - $^{13}\text{C}$  couplings can be observed for RFDR, PDS and DARR (Figures 1.3b-d). With increasing mixing times, two-bond, three-bond and even more distant coherence transfers between dipolar-coupled nuclei occur. In accordance with Zech *et al.* (2005),<sup>43</sup> DARR performs better than PDS (indicated by stars in Figure 1.3d). The RF field applied during mixing on the proton channel in the DARR pulse sequence assists the magnetization transfer between dipolar-coupled carbons by  $^1\text{H}$ - $^{13}\text{C}$ -recoupling. Nevertheless, regardless of the pulse sequence low mixing times were sufficient to almost completely assign 1D  $^{13}\text{C}$  CP resonances. Signals belonging to carbonyl carbons (CO) are generally shifted downfield,

while aliphatic carbons can be found upfield different from aromatic carbons enabling a preliminary classification of carbon resonances. In peptides and proteins, each carbonyl is connected to its own C $\alpha$  and to the backbone nitrogen of the sequential amino acid except for the C-terminus (the formyl carbonyl of N-f-MLF-OH is connected to the methionine nitrogen only and could be detected as low-intensity signal due to missing  $^{13}\text{C}$ -labelling). Cross peaks in the carbonyl region observable in spectra of low mixing times thus could be assigned to correlations between CO and C $\alpha$  atoms. Knowing C $\alpha$  chemical shifts, remaining signals could be allocated to side chain carbons using the same logic. Site-specific assignment was achieved by structural differences between the three amino acids methionine, leucine and phenylalanine, i.e. the C $\gamma$  atom of phenylalanine is part of the aromatic ring and methionine contains a sulfur atom at the  $\delta$ -position.



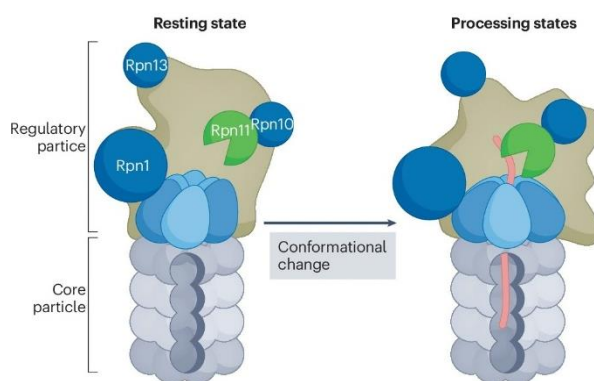
**Figure 1.3:** (a) Structure of the chemotactic tripeptide N-formyl-L-Met-L-Leu-L-Phe (N-f-MLF-OH). (b)  $^{13}\text{C}$ - $^{13}\text{C}$  RFDR spectrum of N-f-MLF-OH with a mixing time of 0.8 ms. (c) and (d)  $^{13}\text{C}$ - $^{13}\text{C}$  PDSD and DARR spectra of N-f-MLF-OH with a mixing time of 5 ms, respectively. Differences are indicated as \* in the DARR spectrum.

In the case of the tripeptide N-f-MLF-OH, 2D  $^{13}\text{C}$ - $^{13}\text{C}$ -correlation spectra not only allow for the identification of spin systems and amino acid types, but also the site-specific assignment of carbon resonances. Proteins contain hundreds of residues and corresponding spectra are crowded. The sequential linking of identified and amino acid typed spin systems in a uniformly labelled protein

thus requires the recording of 3D and high-dimensional ( $\geq 4D$ ) spectra. One common method for the site-specific assignment of carbon and nitrogen resonances includes the separate acquisition of 3D  $^{15}N$ - $^{13}C$ - $^{13}C$  NCOCX and NCACX (CX = any carbon) spectra using SPECIFIC-CP<sup>44</sup> from backbone nitrogen atoms either to CO or C $\alpha$  carbons and subsequent carbon mixing.<sup>45</sup> Experiments such as ZF TEDOR-DARR<sup>46</sup>/RFDR<sup>47</sup> feature a bi-directional polarization transfer enabling the simultaneous acquisition of NCOCX and NCACX connectivities within one spectrum. The pulse sequence of ZF TEDOR-DARR is shown in Figure 1.2b.

### 1.3 Proteasomal degradation pathways in eukaryotic cells

The 26S proteasome is a multi-protein complex in eukaryotic cells. It is responsible for the ATP-dependent degradation of virtually all cellular proteins to maintain proteostasis<sup>48</sup> (protein homeostasis), i.e. a balanced proteome. The generated peptides of this process find use in cell signalling and antigen presentation, are recycled for the synthesis of new proteins,<sup>49</sup> or have anti-microbial properties<sup>50</sup>. The 26S proteasome is constituted by the barrel-shaped 20S core particle and the 19S regulatory particle as shown in Figure 1.4. The core particle contains four stacked heptameric rings (two outer  $\alpha$ -rings enclosing two inner  $\beta$ -rings), which form the degradation chamber. Three of the seven  $\beta$ -subunits have proteolytic active sites, whereas the  $\alpha$ -subunits function as gate to control substrate access to the degradation chamber. The regulatory particle is built of the lid of eight scaffolding subunits plus the deubiquitinase Rpn11 and the base. The base contains the scaffolding subunit Rpn2, the receptor subunits Rpn1, Rpn10 and Rpn13 as well as the AAA+ motor with six ATPase subunits (Rpt1-6). Recruited substrates can enter the central channel of the hexameric ATPase motor for engagement leading to a conformational switch of the proteasome (Figure 1.4). In the processing state, the core particle gate is opened and ATP-hydrolysis drives the unfolding and translocation of the substrate into the proteolytic chamber.<sup>49</sup>



**Figure 1.4:** The 26S proteasome consists of the 19S regulatory particle and the 20S core particle. Upon substrate engagement, the proteasome changes conformation from the resting into the processing state. Taken from ref. [49].

Recruited substrates can enter the central channel of the hexameric ATPase motor for engagement leading to a conformational switch of the proteasome (Figure 1.4). In the processing state, the core particle gate is opened and ATP-hydrolysis drives the unfolding and translocation of the substrate into the proteolytic chamber.<sup>49</sup>

Posttranslational modifications (PTM) like phosphorylation or glycosylation are a common way to regulate the function, activity or localization of proteins. Other than small molecules, whole proteins can function as PTM: Upon the covalent attachment of multiple ubiquitin moieties, substrate proteins are targeted for degradation by the 26S proteasome. The heat-stable protein of 8.6 kDa is characterized by a mixed secondary structure known as the  $\beta$ -grasp fold and carries a diglycine motif at its very C-terminus for conjugation by a conserved enzyme cascade:<sup>51-56</sup> In a first step, ubiquitin is bound in an ATP-dependent manner to the active site cysteine of ubiquitin-activating E1 enzymes, followed by transthioesterification to ubiquitin-conjugation E2 enzymes and eventually, covalent linkage at the C-terminus to substrate proteins catalysed by substrate-

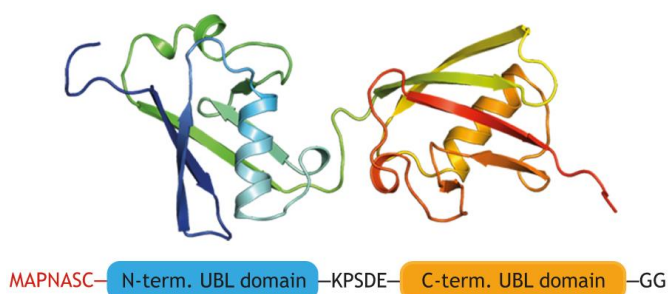
specific E3 ligases.<sup>57</sup> Additional conjugation reactions result in multiple mono-ubiquitinated sites or polyubiquitin chains, which are bound by proteasomal receptors.<sup>49,55,58</sup> Upon the action of deubiquitinases, ubiquitin gets recycled and is available as protein tag in the cell.<sup>59</sup>

To target substrates definitely for degradation by the 26S proteasome, a second signal is required in addition to ubiquitination: Unstructured initiation regions of sufficient length and complexity have to be present within the substrate proteins for insertion and engagement with the ATPase motor.<sup>49</sup> Complexity in the sense that the amino acid composition, structure and dynamics of the initiation region as well as the surrounding local structure and the relative position of the ubiquitin modification are observed to have an effect on the probability of unfolding and the degradation kinetics. For example, the efficiency of unfolding depends on the stability and organization of the local structure following the flexible initiation region and thus, it may be harder to pull a  $\beta$ -strand apart from a  $\beta$ -sheet than unwinding a helix or a loop. In the absence of an unstructured initiation region, pre-processing by the segregase VCP (valosin containing protein, also known as p97) still can lead to ubiquitin-dependent degradation of even tightly folded substrate proteins.<sup>49</sup>

Proteasomal degradation is a complex process and (poly)ubiquitination cannot be thought as the universal tag to dedicate substrates for degradation.<sup>60</sup> Actually, alternative proteasomal degradation pathways independent of ubiquitin are not exceptional cases and just a few can be mentioned here. The existence of ubiquitin-independent proteasomal degradation enables regulation on different levels. Oversimplified, as long as proteins show an unstructured region of sufficient length and have some ways and means to associate with proteasomes, they can be degraded.<sup>61</sup>

Ornithine decarboxylase (ODC) is one of the cellular proteins that gets degraded independently of ubiquitin. The enzyme is involved in polyamine biosynthesis and forms a homodimer in its active form. The protein antizyme 1 binds ODC monomers preventing dimerization, thereby triggering ubiquitin-independent proteasomal degradation as the C-terminal unstructured tail of ODC gets accessible.<sup>49,61</sup> Another degradation mechanism without the needs for ubiquitination is represented by the midnolin-proteasome pathway. The shuttle factor midnolin facilitates the proteasomal degradation of many nuclear proteins. It consists of three domains, namely an N-terminal ubiquitin-like domain, a C-terminal  $\alpha$ -helix and a "Catch" domain. The Catch domain binds unstructured regions within substrates that show a propensity to form  $\beta$ -strands.<sup>62</sup> At the site of the 26S proteasome, midnolin interacts with the subunit Rpn1 as well as with Rpn11. The deubiquitinase Rpn11, which normally removes ubiquitin modifications from substrates, functions as a non-enzymatic receptor for midnolin to position bound substrates directly above the degradation channel.<sup>63</sup>

Finally, the ubiquitin-like modifier (ULM) FAT10 is mentioned as initiator of ubiquitin-independent proteasomal degradation. Human leukocyte antigen-F adjacent transcript 10 (FAT10) comprises two domains (N- and C-domain), which are connected by a flexible linker. Both domains show the typical ubiquitin  $\beta$ -grasp fold and the protein bears a C-terminal diglycine motif (Figure 1.5).<sup>55,64</sup> The conjugation cascade is comparable to ubiquitin and includes activation of FAT10 at the C-terminal glycine by the E1 enzyme UBA6. Subsequently, the E2 enzyme USE1 gets charged by



**Figure 1.5:** Structure of the ubiquitin-like modifier FAT10. The N-domain is shown in blue and the C-domain in orange together with the flexible linker KPSDE, the unstructured heptapeptide at the N-terminus and the C-terminal diglycine motif. Taken from ref. [55].

it seems that FAT10 is degraded along with its substrates. Consistently, no processing proteases specific for FAT10 could be identified in cells.<sup>69</sup> In contrast to ubiquitin, degradation of tightly folded substrate proteins is independent of the unfoldase VCP, mainly attributed to the loose and flexible nature of FAT10 compensating for a missing substrate initiation region.<sup>70,71</sup> FAT10 interacts with the 26S proteasome either binding directly to the subunit Rpn10 or indirectly *via* NEDD8 ultimate buster-1 long (NUB1L). With its C-terminal ubiquitin-like (UBL) domain, FAT10 docks to the VWA domain of Rpn10, while the N-terminal UBL domain is used in binding to NUB1L.<sup>72-74</sup> FAT10 and consequently its conjugates fulfil the requirements for proteasomal degradation. Still, the exact mechanism of how the weak folding of FAT10 enables rapid and direct degradation and the role of its interaction with NUB1L need to be determined.

UBA6 with FAT10 *via* a thioester bond. Both, UBA6 and USE1 are bispecific enzymes in the sense of activating and conjugating FAT10 and ubiquitin, whereas covalent auto-modification of USE1 can only be observed for FAT10 and not for ubiquitin. E3 ligases like Parkin catalyse the last step covalently linking FAT10 to substrate proteins.<sup>65-68</sup> Mono-FAT10ylation is sufficient for degradation by the 26S proteasome and

### 1.4 Thesis outline

Chapters 2 and 3 present the first MAS NMR study of FAT10 and its interaction with the adapter protein NUB1L. The ubiquitin-like modifier FAT10 is best known for its function as proteasomal degradation tag.<sup>69</sup> In contrast to ubiquitin, FAT10 is not recycled<sup>69</sup> and shows a weak and loose folding<sup>70</sup>. Concomitantly, the protein tends to aggregate complicating any structural investigation.<sup>55,70,75</sup> Theng *et al.* (2014) eventually succeeded in the determination of the structure of the FAT10 N-domain, by solution-state NMR, upon the deletion of seven N-terminal residues (PDB-ID 2MBE).<sup>75</sup> Later, Aichele *et al.* (2018) could show that mutations of the cysteine residues stabilize FAT10.<sup>70</sup> In this way, the cysteine-free (Cys-free) N-domain could be crystallized and structurally characterized by X-ray crystallography (PDB-ID 6GF1). Crystallization was not possible for the Cys-free C-domain, but structural insights could be gained by solution-state NMR (PDB-ID 6GF2). Stabilization of FAT10 does not interfere with its conjugation to substrate proteins in cells. However, proteasomal degradation of Cys-free FAT10 monomers and respective conjugates is decelerated in comparison to wild-type FAT10 and conjugates.<sup>70</sup> Further structural characterization is critical to understand how the flexible and weakly folded nature of FAT10 results in rapid and direct degradation by the 26S proteasome.

In Chapter 2, the interaction of the Cys-free N-domain of FAT10 with NUB1L is investigated by MAS NMR spectroscopy. Sequential assignments and empirical predictions of backbone torsion angles and secondary structure show that a largely unfolded N-domain of FAT10 is bound by the formation of an intermolecular  $\beta$ -sheet and docking of the N-terminal residue to NUB1L. This suggests that the unfolded residue chain of the FAT10 N-domain is presented by NUB1L to the proteasome, where it functions as initiation region for rapid degradation.

Chapter 3 reports the wild-type N-domain of FAT10 as studied by MAS NMR spectroscopy. With increasing concentrations, the protein starts to aggregate and corresponding spectra are of poor quality. Upon co-sedimentation with NUB1L, the resolution improves tremendously. This proposes co-sedimentation with a known binding partner as general strategy for the MAS NMR sample preparation of weakly folded proteins. NUB1L binding of the wild-type N-domain is identical to the Cys-free N-domain and involves a stretch of residues that become part of an intermolecular  $\beta$ -sheet and docking of the N-terminus.

## 2 Chapter 2 Intermolecular $\beta$ -sheet formation guides interaction between FAT10 and NUB1L

### 2.1 Abstract

Under inflammatory conditions, the ubiquitin-like modifier FAT10 targets proteins for rapid and irreversible degradation by the 26S proteasome. FAT10 is degraded along with its substrates; in this process, the weak folding of FAT10 and the adapter protein NUB1L are believed to play crucial roles. We report the investigation of the N-domain of FAT10 and its interaction with NUB1L by magic-angle spinning (MAS) nuclear magnetic resonance (NMR) spectroscopy. A stretch of residues that is unobserved when the N-domain of FAT10 is in its ubiquitin-like  $\beta$ -grasp fold, becomes part of an intermolecular  $\beta$ -sheet upon binding to NUB1L. The rest of the N-domain is stabilized in a disordered state, with exception of a series of anchor residues and the N-terminus. MAS NMR is thus able to experimentally capture, at the atomic level, the shapeshifting of FAT10. Binding of the N-terminus likely plays a critical role in the presentation of the unfolded N-domain to the proteasome.

### 2.2 Introduction

The protein FAT10 (human leukocyte antigen-F adjacent transcript 10) consists of two ubiquitin-like (UBL) domains, often referred to as the N- and the C-domain. Its gene was first identified in 1996 by sequencing of the human major histocompatibility complex class I locus,<sup>64</sup> suggesting a role in the immune response. Indeed, FAT10, which is unique to mammals, is mainly present in cells and organs of the immune system.<sup>76-78</sup> In other cell types and tissues, expression of FAT10 is strongly upregulated by the proinflammatory cytokines tumor necrosis factor- $\alpha$  and interferon- $\gamma$ .<sup>79,80</sup> FAT10 plays a role in adaptive as well as innate immune responses, affecting, for example, peptide presentation<sup>81</sup> and autophagy<sup>82</sup> — the list of confirmed covalent and non-covalent binding partners continues to grow.<sup>55</sup> FAT10 has been shown to be upregulated in at least a dozen types of cancer and appears to have a positive impact on metastasis.<sup>77,83</sup> Conversely, disruption of the interaction of FAT10 with MAD2 (mitotic arrest deficient 2) has been shown to inhibit tumor progression.<sup>75</sup> While many of the functions of FAT10 have been investigated only superficially, it is well established that FAT10 targets proteins for degradation by the 26S proteasome.<sup>69</sup> Just like ubiquitylation, FAT10ylation requires a conjugation cascade involving three enzyme types: E1 activating enzyme UBA6<sup>65,66</sup>, E2 conjugating enzyme USE1<sup>67</sup> (and several others)<sup>84</sup>, and E3 ligase Parkin<sup>68</sup>. FAT10 is capable of activating all proteolytic activities of the proteasome, but only together with its binding partner NUB1L, the long isoform of NEDD8-ultimate buster 1.<sup>72,85</sup> Recently, it was found that interaction with the N-domain of FAT10 increases the affinity of NUB1L for the RPN1 subunit of the 19S regulatory particle of the proteasome.<sup>86</sup> Unlike ubiquitin, FAT10 is degraded along with its substrates and, with a half-life in cells of 1 h, it is extremely short-lived.<sup>69</sup> In combination with fast upregulation, this makes FAT10 a tool for prompt but time-limited responses to harmful factors during inflammatory processes.<sup>55</sup> The weak folding of FAT10 has long been suspected to be critical for its biological function. The segregase VCP (valosin containing

protein, also known as p97) is required for proteasomal degradation of ubiquitylated proteins that lack a disordered initiation region, but degradation of FAT10 or FAT10ylated proteins is VCP independent.<sup>70</sup> FAT10 melts at a temperature of 41 °C (compared to 83 °C for ubiquitin)<sup>70</sup> and shows poor solubility<sup>87,88</sup>. The latter property has interfered with structural investigations over the years. Theng *et al.* eventually determined, by liquid-state NMR, a structure of the N-domain after deletion of seven N-terminal residues.<sup>75</sup> Later, Aichele *et al.* succeeded in determining the structure of a stabilized, Cys-free mutant of FAT10 by X-ray crystallography (for the N-domain) and liquid-state NMR (for the C-domain).<sup>70</sup> Both domains show the  $\beta$ -grasp fold of ubiquitin<sup>54</sup>, but their surfaces differ from ubiquitin and each other. The Cys-free mutant was found to be well conjugated, but, confirming suspicions, was degraded at a much slower pace. This raises the question of how the weak folding of FAT10 enables fast degradation and why interaction with NUB1L is essential.

We report here the investigation of the N-domain of FAT10 and its interaction with NUB1L by magic-angle spinning (MAS) NMR spectroscopy. MAS NMR is not restricted by molecular size, solubility or a disorderly arrangement of structural units and can, in addition to inter-atomic distances, probe local chemical environments and motion. This makes MAS NMR well-suited for the atomic-level exploration of interactions between biomolecules, even if one or more binding partners are weakly folded. Recent examples include structural characterization of branching stimulant TPX2 (targeting protein for Xklp2), which forms condensates upon binding to microtubules, but is disordered in solution<sup>89</sup> and the observation of a paired helical filament fold of the intrinsically disordered tau protein bound to lipid membranes<sup>90</sup>. Co-sedimentation by ultracentrifugation is often a good strategy to form biomolecular complexes in the MAS rotor.<sup>91,92</sup>

We start by examining the isolated N-domain of Cys-free FAT10 (N-FAT10-C0) in microcrystalline form. Sequential assignment of resonances is possible for 69 out of 83 residues; not observed are the flexible start and end tails and residues I68-I74. Torsion angles and secondary structure predictions based on chemical shifts are in agreement with the crystal structure<sup>70</sup>. Next, we examine, as a physiologically relevant reference sample, isolated lyophilized and rehydrated N-FAT10-C0. Backbone connectivity is reduced, but with help from the chemical shifts of the microcrystalline sample, sequential assignment remains possible. Predicted torsion angles are similar to the microcrystalline form, but the stretch of unobserved residues is lengthened to I68-L76. Finally, we turn to N-FAT10-C0 co-sedimented with NUB1L. The spectra show a drastically reduced number of resonances, which are confidently assigned to residues I68-V81 and the N-terminus of N-FAT10-C0. We attribute the disappearance of the other residues to disorder. Analysis of chemical shifts indicates that residues T73-V81 of N-FAT10-C0 constitute a  $\beta$ -strand, which provides experimental evidence for AlphaFold-Multimer<sup>93</sup> predictions that suggest the formation of an intermolecular, anti-parallel  $\beta$ -sheet consisting of residues L76-V81 of FAT10 and Y212-N217 of NUB1L. AlphaFold-Multimer, however, does not predict the binding of the N-terminus.

### 2.3 Materials and methods

Below we give short accounts of all procedures. Further specifics can be found in the Supporting Information.

**Expression and Purification of U-<sup>13</sup>C, <sup>15</sup>N-N-FAT10-C0.** Cys-free N-domain of human FAT10 (amino acids 5-86; C7T, C9T) was expressed as a His<sub>6</sub>-GST-fusion protein in *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells (Agilent Technologies). An additional glycine residue at the N-terminus remains from the TEV protease cleavage site. For uniform <sup>13</sup>C and <sup>15</sup>N labelling, cells were grown in M9 minimal medium supplemented with U-<sup>13</sup>C<sub>6</sub>-D-glucose and <sup>15</sup>N-ammonium chloride. The purification was adapted from a previously published protocol.<sup>70</sup> Briefly, bacterial cells were lysed and the supernatant was applied to nickel affinity chromatography. After buffer exchange and cleavage of the purification tag, His-tagged TEV protease and byproducts were separated by a second nickel affinity chromatography. Final purification by size-exclusion chromatography yielded up to 7 mg per liter of M9 medium.

**Expression and Purification of NUB1L.** Human NUB1L (amino acids 2-615) was expressed as a His<sub>6</sub>-SUMO-fusion protein, also in *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells. To produce natural abundance NUB1L, cells were grown in LB medium. Purification requires the same steps as for U-<sup>13</sup>C, <sup>15</sup>N-N-FAT10-C0. Yield was up to 15 mg per liter of LB medium.

**Microcrystalline N-FAT10-C0.** The crystallization protocol of Aichele *et al.*<sup>70</sup> was taken as a starting point to obtain microcrystals of U-<sup>13</sup>C, <sup>15</sup>N-N-FAT10-C0. The sitting drop vapor diffusion method was used, with a crystallization solution of ammonium sulfate at pH 2.5. Microcrystals were packed into the MAS rotor by ultracentrifugation with a tool manufactured at the workshop of the University of Konstanz following Mandal *et al.*<sup>94</sup> The rotor finally contained an estimated 13.5 mg of protein.

**Lyophilized/rehydrated N-FAT10-C0.** With its molecular weight of less than 10 kDa, the N-domain of FAT10 is too small to be sedimented into a MAS rotor by ultracentrifugation.<sup>18</sup> To nevertheless obtain a reference sample at physiological pH, U-<sup>13</sup>C, <sup>15</sup>N-N-FAT10-C0 was lyophilized from buffer solution (pH 7.5), ground, and manually packed into a rotor, to a total estimated amount of 14.5 mg. To rehydrate, MilliQ water was added in steps of 1  $\mu$ L until spectral resolution no longer improved.

**Complex of N-FAT10-C0 and NUB1L.** Binding to NUB1L was previously demonstrated for the wild-type N-domain of FAT10<sup>74</sup>, but not for the stabilized, Cys-free version. We therefore verified complex formation of N-FAT10-C0 and NUB1L by size exclusion chromatography (Figure S2.1). Co-sedimentation by ultracentrifugation, in a molar ratio of 1:1, was used to form the complex of natural abundance NUB1L and U-<sup>13</sup>C, <sup>15</sup>N-N-FAT10-C0 in the MAS rotor.

**MAS NMR Spectroscopy.** Experiments were performed at 18.8 and 20.0 T (<sup>1</sup>H Larmor frequencies of 800 and 850 MHz) on Bruker Avance NEO/III spectrometers, each equipped with a 3.2 mm E-free HCN Bruker MAS probe. The spinning frequency was 14.5 or 19.0 kHz. The sample temperature was 4 °C, unless noted otherwise. Chemical shifts of <sup>13</sup>C are referenced to DSS in D<sub>2</sub>O (0.5 % by weight), chemical shifts of <sup>15</sup>N are referenced to liquid ammonia at 25 °C.

Sample integrity was verified by <sup>1</sup>H-<sup>13</sup>C and <sup>1</sup>H-<sup>15</sup>N cross-polarization experiments as well as <sup>13</sup>C-<sup>13</sup>C correlation spectra recorded with the dipolar assisted rotational resonance (DARR) pulse sequence.<sup>34</sup> The mixing time was set to 10.0 ms to observe primarily one- and two-bond correlations. From <sup>13</sup>C-<sup>13</sup>C correlation patterns, amino acid residue types can be identified, but sequential assignment is not possible. For this purpose, <sup>15</sup>N-<sup>13</sup>C and <sup>15</sup>N-<sup>13</sup>C-<sup>13</sup>C spectra were recorded with the z-filtered transferred

echo double-resonance (ZF TEDOR)<sup>32</sup> and the ZF TEDOR-DARR<sup>46</sup> pulse sequences, respectively. The TEDOR mixing time was set to ~1 ms to emphasize N-C $\alpha$  and N-C' magnetization transfers along the protein backbone and, in the side chains, to carbon nuclei directly bonded to nitrogen nuclei. The DARR mixing time in these experiments was set to 40-50 ms to enable transfer from the C $\alpha$  and C' nuclei to other carbon nuclei within a residue.

The ZF TEDOR-DARR sequence provides NCACX and NCOCX connectivities within one 3D spectrum, but at the cost of large spectral widths. To keep the number of increments in the indirect dimensions under control, it sufficed for microcrystalline N-FAT10-C0 to truncate the free induction decay along the indirect <sup>13</sup>C dimension and pad it using linear forward prediction. For lyophilized/rehydrated N-FAT10-C0 and the N-FAT10-C0/NUB1L complex, we used aliasing. At a spinning frequency of 19 kHz, a convenient strategy is to set the dwell time of the indirect <sup>13</sup>C dimension equal to one rotor period and the carrier frequency to approximately 50 ppm.<sup>47</sup> The carbonyl resonances are now observed at the positions of their spinning sidebands just above the aliphatic region. A supplementary NCOCX measurement with SPECIFIC-CP<sup>44</sup> and combined R2<sub>n</sub><sup>17</sup>-driven (CORD)<sup>95</sup> <sup>13</sup>C-<sup>13</sup>C mixing provided additional connections for lyophilized/rehydrated N-FAT10-C0.

An overview of all MAS NMR experiments and acquisition parameters is given in Tables S2.1 and S2.2. Spectra were processed in TopSpin or NMRPipe<sup>96</sup>. CcpNmr Analysis was used for the resonance assignments.<sup>17,97</sup>

**Torsion Angle and Secondary Structure Prediction.** Backbone torsion angles  $\Phi$  and  $\Psi$  and secondary structure were predicted empirically based on chemical shifts of assigned N, C', C $\alpha$ , and C $\beta$  nuclei.<sup>98,99</sup> For this purpose, we used TALOS-N,<sup>100</sup> which combines a set of trained neural networks with efficient mining of a database of proteins of which both the X-ray structure and the chemical shifts are known.

**AlphaFold Modelling.** The structure of isolated N-FAT10-C0 (amino acids 5-86 with C7T and C9T and the additional glycine at the N-terminus) was predicted with AlphaFold<sup>8</sup> v2.3.2 (reduced database). The maximum template release date was 2014-08-26. Five predictions were obtained, from one seed per model, and ranked according to the predicted local distance difference test (pLDDT) confidence score; only the best prediction was relaxed.

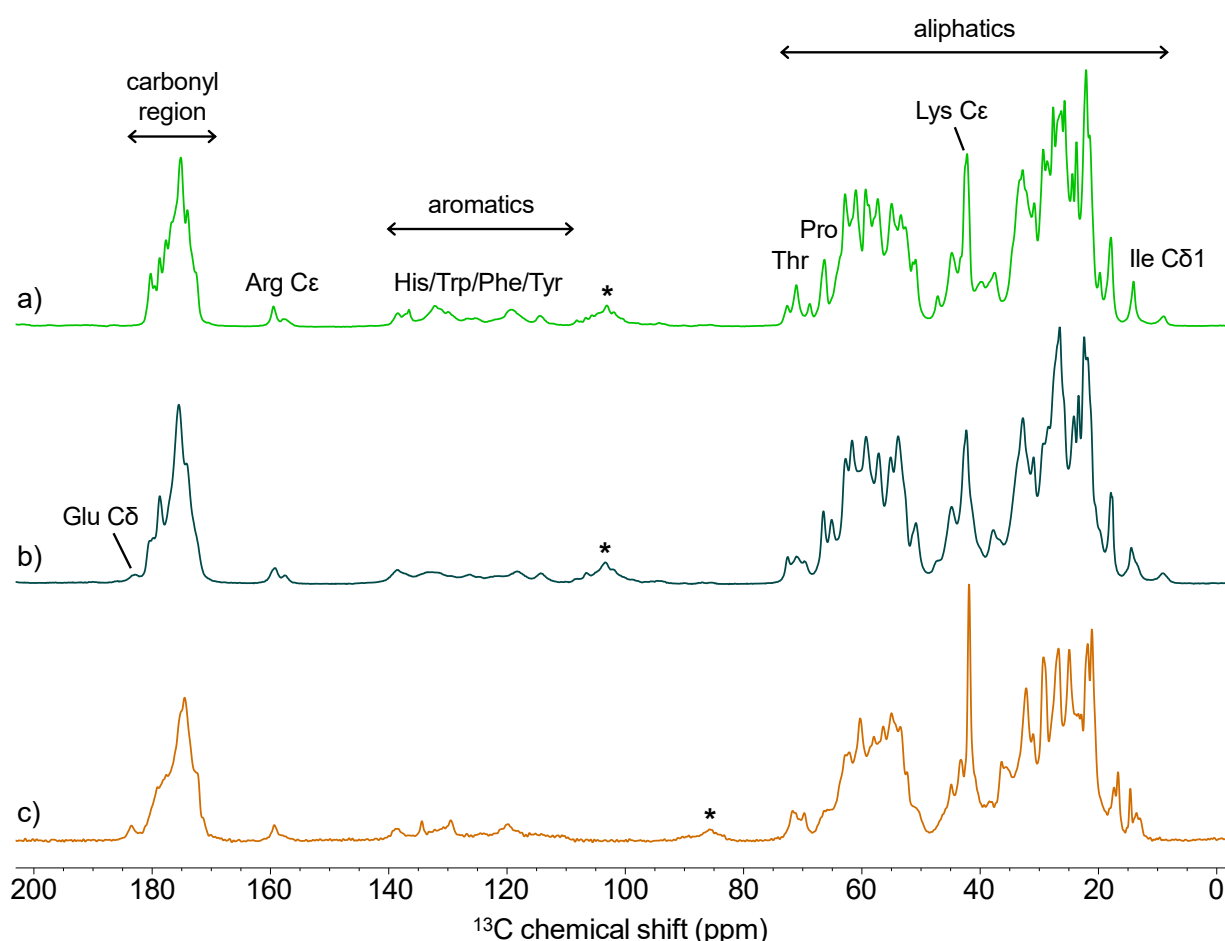
The structure of the complex of N-FAT10-C0 and NUB1L (amino acids 1-615) was predicted with AlphaFold-Multimer (reduced database)<sup>93</sup>. The maximum template release date was 2004-11-28. Twenty-five predictions were obtained, from five seeds per model, and ranked according to a weighted combination of the predicted template modelling (pTM) and interface pTM scores (model confidence = 0.8·ipTM + 0.2·pTM). Only the best prediction was relaxed.

All AlphaFold modelling was performed on the Scientific Compute Cluster of the University of Konstanz (SCCKN). UCSF ChimeraX<sup>101</sup> was used for structure visualization. Torsion angles were extracted from the AlphaFold structures using VADAR<sup>102</sup>.

## 2.4 Results

### 2.4.1 Valuation of Samples

Figures 2.1 and S2.2 show the  $^1\text{H}$ - $^{13}\text{C}$  and  $^1\text{H}$ - $^{15}\text{N}$  cross-polarization spectra of the three samples of N-FAT10-C0. Conformational homogeneity of N-FAT10-C0 in the microcrystals<sup>103,104</sup> leads to well-resolved spectra (bright green);  $^{13}\text{C}$  linewidths are 1 ppm or less. Some resonances are skewed, for example, of  $^{13}\text{C}_{\delta 1}$  of Ile at 9 ppm. This likely reflects distinct conformations of multiple molecules in the unit cell. In the sample of lyophilized/rehydrated N-FAT10-C0 (turquoise spectra), the conformations of the protein are not restricted by crystal packing and resonances have broadened; this is particularly noticeable for the backbone nitrogen atoms. Deprotonation of Asp and Glu has shifted the  $^{13}\text{C}_{\gamma}$  and  $^{13}\text{C}_{\delta}$  resonances downfield,<sup>105,106</sup> changing the appearance of the carbonyl region; for the same reason, the resonance of  $^{15}\text{N}_{\delta 1}$  of His has shifted downfield, out of the spectral range. Although there are further small changes, the  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization spectra of lyophilized/rehydrated N-FAT10-C0 and of microcrystalline N-FAT10-C0 look very similar. To prepare the third sample, N-FAT10-C0 was co-sedimented with NUB1L. Resolution of the spectra (orange) is good, with  $^{13}\text{C}$  linewidths again  $\leq 1$  ppm. Spectral features differ drastically from isolated N-FAT10-C0.



**Figure 2.1:**  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization spectra of (a) microcrystalline N-FAT10-C0, (b) lyophilized/rehydrated N-FAT10-C0, and (c) N-FAT10-C0 in complex with NUB1L. Spectra were recorded at 18.8 T with a spinning frequency of 14.5 kHz (a,b) and at 20.0 T with a spinning frequency of 19.0 kHz (c). The sample temperature was 4 °C. The stars indicate spinning sidebands.

## 2.4.2 Resonance Assignments and Secondary Structure of Microcrystalline N-FAT10-C0

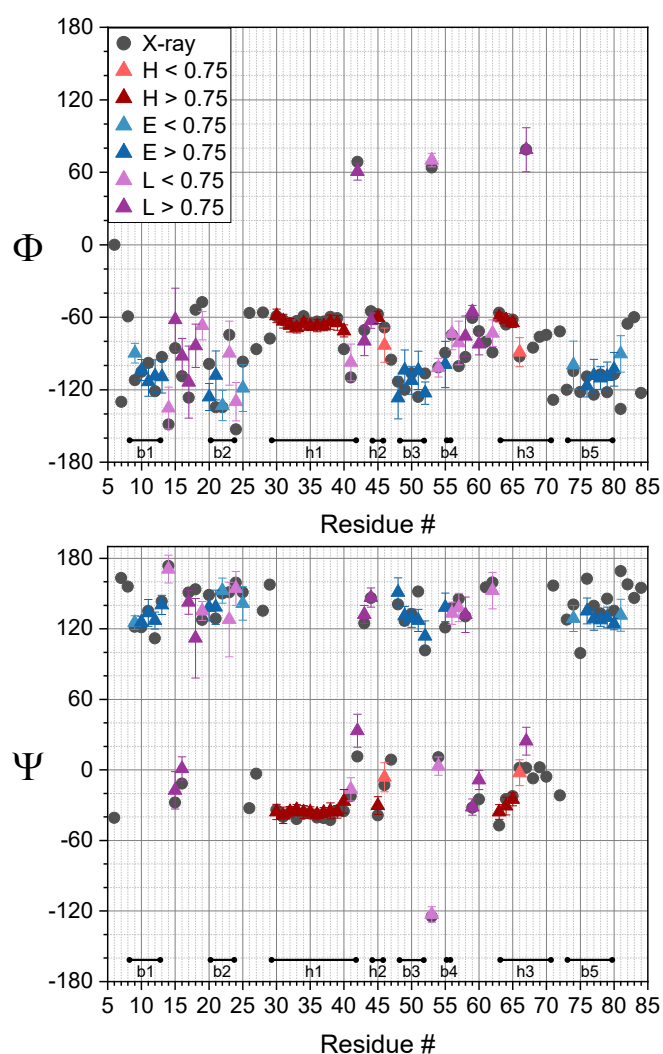
$^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectra of microcrystalline N-FAT10-C0 are shown, along with resonance assignments, in Figures S2.3a and b. Figure S2.4 shows a representative strip plot of the  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  spectrum. Connectivity along the backbone is excellent: following the NCA/NCO transfer, side chains are visible up to the  $^{13}\text{C}_\alpha$ . In tightly packed regions, magnetization also transfers to  $^{13}\text{C}$  nuclei of neighbouring residues, which was helpful to verify sequential assignments. The  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectrum is dominated by one- and two-bond transfers, but cross peaks from three-bond transfers are also discernible from the noise. This made for convenient pattern recognition and assisted in the assignment of residues with aromatic carbons or long aliphatic side chains. Some inter-residue contacts, e.g. T9-V10 and G53-S54, are visible. The  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectrum served to verify connectivities observed in the 3D spectrum and was helpful for the assignment of residues with a  $^{15}\text{N}$  nucleus in the side chain.

In total, we were able to sequentially assign 69 of the 83 residues of N-FAT10-C0, see Figure 2.2. For five of these residues, some uncertainty remains (highlighted yellow in Figure 2.2). The backbone  $^{15}\text{N}$  chemical shifts of K32 and K34 and their succeeding residues are nearly identical, making it difficult to tell them apart. The signals from L8, D69, and T73 are very weak. For eight residues, we observed up to three conformations (blue). All unobserved residues (grey) are located in the flexible tails (G4-T7 and K82-E86) and in the stretch I68-I74.

Microcryst: GASTLTVHVRSEEWDLMTFDANPYDSVKKIK<sup>Y</sup>EHVRSKTKVPVQDQVLLGSKILK<sup>Y</sup>PRRSLSSYGIDKEKTIHLTLK<sup>Y</sup>VVKPSDE  
 Lyoph/rehydr: GASTLTVHVRSEEWDLMTFDAN<sup>Y</sup>PYDSVKKIK<sup>Y</sup>EHVRSKTKVPVQDQVLLGSKILK<sup>Y</sup>PRRSLSSYGIDKEKTIHLTLK<sup>Y</sup>VVKPSDE  
 w NUB1L: GASTLTVHVRSEEWDLMTFDANPYDSVKKIK<sup>Y</sup>EHVRSKTKVPVQDQVLLGSKILK<sup>Y</sup>PRRSLSSYGIDKEKTIHLTLK<sup>Y</sup>VVKPSDE  
           5    10    15    20    25    30    35    40    45    50    55    60    65    70    75    80    85

**Figure 2.2:** Overview of sequential assignments of microcrystalline N-FAT10-C0, lyophilized/rehydrated N-FAT10-C0, and N-FAT10-C0 in complex with NUB1L. Green = unambiguous, yellow = ambiguous, blue = multiple conformations, grey = unobserved. Of the grey, underlined residues of N-FAT10-C0 in complex with NUB1L, we observe 1×E, 1×I, 1×L, 1×P, 1×R, 2×S, and 3×V, but sequential assignment was not possible.

Figure 2.3 shows the predicted backbone torsion angles  $\Phi$  and  $\Psi$  based on the resonance assignments, together with the torsion angles of the X-ray structure of N-FAT10-C0<sup>70</sup>. Agreement is very good; the root of the mean square deviation is 14° for  $\Phi$  and 12° for  $\Psi$ . Predicted secondary structure elements also match the X-ray structure (Figure S2.3c).



**Figure 2.3:** Backbone torsion angle and secondary structure predictions based on MAS NMR of microcrystalline N-FAT10-C0. Error bars correspond to the standard deviations of the  $\Phi$  and  $\Psi$  angles of the best matches in the TALOS-N database. For residues P26-S29, D47, R61, and H75, there is no consensus among the database matches and, hence, no torsion angles are shown. No predictions are made for L8, I68-T73, and K82. Secondary structure predictions are based on observed chemical shifts except for G67, for which the prediction is based on sequence information, and classified as helix (H), strand (E), or loop (L). Predictions with a confidence > 0.75 are highly reliable. Torsion angles (black dots) and secondary structure elements (b stands for strand and h for helix) of the X-ray structure of N-FAT10-C0 (PDB-ID 6GF1, chain B) are also plotted for comparison.

### 2.4.3 Resonance Assignments and Secondary Structure of Lyophilized/Rehydrated N-FAT10-C0

$^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectra of lyophilized/rehydrated N-FAT10-C0 are shown, along with resonance assignments, in Figures 2.4a and b. As expected from the one-dimensional spectra, cross peaks have broadened, but, in the  $^{13}\text{C}$ - $^{13}\text{C}$  spectrum, side chain patterns are still readily recognized. Figure S2.5 shows a representative strip plot of the  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  spectrum. Backbone connectivity is not as good as for the microcrystalline sample. This problem was mitigated by recording an additional NCOX correlation spectrum (Figure S2.6), with a narrow spectral width and an increased number of scans.

In total, we sequentially assigned 56 of the 83 residues of N-FAT10-C0 (Figure 2.2). The assignment was performed *de novo*, but the known chemical shifts of microcrystalline N-FAT10-C0 provided essential guidance. For example, the side chain patterns of I33 and I56 are very similar in both samples. For residues T21, P26, Y27, L51, L63, and L78, the connectivities were only observed in the additional NCOCX spectrum. Signals from P26 and P56 partially overlap and signals from their preceding residues (N25 and K58) are weak. Hence, these assignments are ambiguous (yellow). The same applies for S64, which overlaps with other serines; the signals from K79 are weak. Residues in the elongated, flexible tails G4-L8 and V80-E86 are not observed, nor is the full stretch I68-L76 (grey), resembling microcrystalline N-FAT10-C0. In addition, residues D18, F22, H36, V45, Q46, and K55 are not observed; possibly multiple conformations of (neighboring) bulky side chains create local disorder in the backbone.

Comparison of the chemical shifts of the microcrystalline and lyophilized/rehydrated samples of N-FAT10-C0 shows small changes (1-2 ppm) for most nuclei, which tend to be larger, up to 5 ppm, if a neighboring residue is not observed in the lyophilized/rehydrated sample. Vice versa, N-C $\alpha$  connectivities of W17 and G67 were not observed for the microcrystalline sample, but are visible in the  $^{15}\text{N}$ - $^{13}\text{C}$  spectrum of the lyophilized/rehydrated sample (Figure 2.4b) at 59.1, 118.8 ppm and 45.9, 107.5 ppm, also suggesting that local changes in structure have occurred. Predicted torsion angles and secondary structure (Figure S2.7) are, however, similar for both samples of isolated N-FAT10-C0, indicating that the overall fold is unchanged. This result is visualized in Figure 2.4c, where the predicted secondary structure is projected on the AlphaFold structure of N-FAT10-C0.

**Figure 2.4 (on the following page):** (a)  $^{13}\text{C}$ - $^{13}\text{C}$  DARR spectrum and (b)  $^{15}\text{N}$ - $^{13}\text{C}$  ZF TEDOR spectrum of lyophilized/rehydrated N-FAT10-C0. Sequentially assigned cross peaks are marked and labelled in both spectra. As a result of the relatively long mixing time and the large number of scans for the given amount of protein, two- and three-bond  $\text{NC}_\beta/\text{NC}_\gamma$  correlations are visible in the  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectrum. Cross peaks from side chains of His ( $^{15}\text{N}$  chemical shift 166 ppm), Arg ( $^{15}\text{N}$  chemical shift 67-90 ppm), and Lys ( $^{15}\text{N}$  chemical shift 35 ppm) are also observed, but not plotted here. (c) Projection of TALOS-N secondary structure predictions per residue on the AlphaFold structure of N-FAT10-C0. H stands for helix, E for strand, and L for loop. D18, F22, H36 and K55 are unassigned, but secondary structure is predicted from the chemical shifts of neighboring residues. For residues N25, K58, P59, S64 and K79, the assignments are ambiguous, but since the secondary structure is predicted from chemical shifts, they are color coded accordingly. See Figure S2.7 for plots of the backbone torsion angles and further specifics.



### 2.4.4 Interaction of N-FAT10-C0 and NUB1L

Figures 2.5a and b show the  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectra of N-FAT10-C0 co-sedimented with NUB1L. Compared to the spectra of isolated N-FAT10-C0 (Figures S2.3a,b and 2.4a,b), the number of cross peaks is much reduced. In the aforementioned MAS NMR pulse sequences, magnetization transfer relies on dipolar couplings, which can be averaged out by molecular motion. To explore if this is happening here, we measured  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization spectra at reduced temperatures of  $-2$ , and  $-10$  °C (Figure S2.8). These spectra, however, show no new signals, only a slight decrease in resolution. We also obtained a  $^{13}\text{C}$  spectrum with the liquid-state NMR sequence INEPT (insensitive nuclei enhanced by polarization transfer)<sup>107</sup>, which requires J-coupling for magnetization transfer and motion on pico- to nanosecond time scales. Weak signals from mobile parts of side chains become visible only after a lot of averaging. We surmise that disorder (which may still be dynamic, see the Discussion) renders a large number of amino acid residues undetectable. For the cross peaks that are visible, resolution is good, as is the connectivity; a representative strip plot of the  $^{15}\text{N}$ - $^{13}\text{C}$  spectrum is shown in Figure S2.9.

Confident sequential assignment is possible for residues I68-V81 (Figure 2.2). Intriguingly, the Gly residue at the N-terminus (G4) is visible at 43.8, 26.8 ppm in the  $^{15}\text{N}$ - $^{13}\text{C}$  spectrum (Figure 2.5b). Side chain patterns are evident for eleven residues (1×E, 1×I, 1×L, 1×P, 1×R, 2×S, 3×V, 1×W), but their backbone connectivities could not be observed. Torsion angle and secondary structure predictions based on the chemical shifts of the assigned residues are plotted in Figure S2.10. Residues D69-E71 form a loop, while residues T73-V81 form a  $\beta$ -strand.

To further explore the interaction between N-FAT10-C0 and NUB1L, we used structure prediction by AlphaFold-Multimer. The best-ranked structure is shown in Figure 2.5c (confidence scores are provided in Figure S2.11). N-FAT10C0 is in an open, partially unfolded state and NUB1L is wrapped around an intermolecular, anti-parallel  $\beta$ -sheet that consists of residues L76-V81 of N-FAT10-C0 and Y212-N217 of NUB1L.

The AlphaFold-Multimer prediction suggests a plausible interpretation of the MAS NMR experimental data. In preparation of rapid degradation by the proteasome, interaction with NUB1L stabilizes N-FAT10-C0 in an unfolded, mostly disordered state. Only the residues visible in the spectra retain a regular structure; those that were sequentially assigned are highlighted in orange in Figure 2.5c. The inset in Figure 2.5c shows the TALOS-N secondary structure predictions projected on the AlphaFold-Multimer structure. MAS NMR thus confirms the formation of the intermolecular  $\beta$ -sheet and suggests that it is, in reality, three residues longer (on the side of N-FAT10-C0, residues T73-V81 participate as opposed to L76-V81). The residues of N-FAT10-C0 preceding the  $\beta$ -strand form a loop until I68. In addition, eleven residues scattered along the N-FAT10-C0 chain plus the Gly N-terminus show conformational homogeneity. We suspect that these residues form anchor points where N-FAT10-C0 interacts non-covalently with NUB1L. Of these, the N-terminus is of particular importance because it is the first residue of the unfolded FAT10 chain to enter the proteasome<sup>69-71</sup>.



**Figure 2.5:** (a)  $^{13}\text{C}$ - $^{13}\text{C}$  DARR spectrum and (b)  $^{15}\text{N}$ - $^{13}\text{C}$  ZF TEDOR spectrum of N-FAT10-C0 co-sedimented with NUB1L. Assigned cross peaks are marked and labelled in both spectra. Cross peaks from side chains of Arg ( $^{15}\text{N}$  chemical shift 67-90 ppm) are also observed but not plotted here. (c) AlphaFold-Multimer prediction for the interaction of N-FAT10-C0 and NUB1L. The residues of N-FAT10-C0 that are observed by MAS NMR are highlighted in orange. The inset shows the secondary structure predictions for residues I68-V81 projected on the AlphaFold-Multimer structure (E stands for strand and L for loop). See Figure S2.10 for the backbone torsion angles and further specifics.

## 2.5 Discussion

The E1 enzyme UBA6 is unusual in the sense that it not only activates FAT10, but also ubiquitin.<sup>65,66</sup> Co-crystallization of UBA6 with Cys-free FAT10 revealed that the binding of the C-domain of FAT10 is analogous to the binding of ubiquitin to the, ubiquitin-only, E1 enzyme UBA1.<sup>108</sup> The N-domain of FAT10 interacts with the three-helix bundle of UBA6, with its  $\beta$ -grasp fold intact. The conjugation cascade thus requires FAT10 in properly folded form. The interaction of N-FAT10 with UBA6 bears similarity with the interaction of another ubiquitin-like modifier NEDD8 (neural precursor cell-expressed developmentally down-regulated protein 8) with NUB1L: both interactions require a UBA domain, are polar, and rely on N51 of NEDD8 or the analogous residue K58 in N-FAT10. With this information in mind, we naively expected N-FAT10-C0 to remain folded upon binding to NUB1L. To assess the feasibility of a MAS NMR study, we first prepared a microcrystalline sample of isolated N-FAT10-C0. Next, aiming to gain insight into the binding mode and interface *via* chemical shift perturbations, we prepared a lyophilized/rehydrated reference sample of N-FAT10-C0 at physiological pH. Sequential assignment of N-FAT10-C0 co-sedimented with NUB1L, however, constituted a plot twist and poignantly showed the shapeshifting of the N-domain of FAT10. This experience illustrates the strength of MAS NMR spectroscopy – the unfolding of N-FAT10-C0 makes co-crystallization with NUB1L unlikely to be successful. In contrast, co-sedimentation in the MAS rotor favors the formation of complexes because complexed molecules are extracted from the bulk solution first.<sup>18</sup>

Earlier this year Arkinson *et al.* reported the investigation of the conformation and solvent accessibility of wild-type FAT10 upon binding with NUB1 (the shorter splice variant of NUB1L) using hydrogen-deuterium exchange detected by mass spectrometry.<sup>71</sup> In the absence of NUB1, several peptides from both the N- and the C-domain show a bimodal distribution, indicating coexisting folded and unfolded states. In the presence of NUB1, peptides throughout the N-domain were exposed, except for the last  $\beta$ 5-strand. Thus, NUB1 traps an unfolded state of the N-domain of FAT10. The coexistence of folded and unfolded states may well explain the troubles encountered in the investigation of the N-domain of FAT10 by liquid-state NMR. Theng *et al.* were only successful after deletion of the first seven N-terminal residues and even then residues P59-T73 could not be detected by  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single-quantum correlation (HSQC); in the structure they obtained, the  $\alpha$ 1-helix is somewhat displaced (PDB-ID 2MBE).<sup>75</sup>

Arkinson *et al.* found their conclusion supported by structure prediction with AlphaFold-Multimer and by site-directed mutagenesis: FAT10 mutants H75A, H75D and H75K all compromised NUB1 binding, H75D and H75K also abolished degradation.<sup>71</sup> Cao *et al.* recently went a step further showing that phosphorylation of T77 interferes with NUB1L binding.<sup>109</sup> Interestingly, Arkinson *et al.* detected no complex formation between full-length Cys-free FAT10 and NUB1, while in our hands complex

formation between N-FAT10-C0 and NUB1L is evident (Figure S2.1). Possibly the absence of the C-domain makes it easier for the N-domain to become inserted into the clasp of NUB1L. Finally, Arkinson *et al.* showed by cryo-electron microscopy that FAT10 binding induces an open conformation of NUB1, allowing its ubiquitin-like (UBL) domain to interact with the RPN1 subunit<sup>73</sup> of the regulatory particle for direct FAT10 delivery. In spite of directed efforts, however, it was not possible to resolve the interaction of NUB1 with FAT10.

Structure-based sequence alignment of FAT10 with ubiquitin reveals two extra residues in the N-domain preceding the  $\beta$ 5-strand (K70, E71), extending the loop. Molecular dynamics simulations indicate loose folding in this region,<sup>70</sup> in agreement with the experimental observations by MAS NMR. Moreover, multiple sequence alignment of mammalian FAT10 proteins reveals a highly conserved surface comprising residues T73, H75, L76, T77, L78, and V80. This region of N-FAT10 thus seems primed for formation of the intermolecular  $\beta$ -sheet with NUB1L. This is reminiscent of the capture of an unstructured  $\beta$ -strand degron by two Catch subdomains in the midnolin pathway of ubiquitin-independent protein degradation.<sup>62</sup> In the case of N-FAT10, the primed region appears to constitute an interaction motif. Titrations in combination with liquid-state NMR have implicated the  $\beta$ 1- and  $\beta$ 5-strands of N-FAT10 in the binding with MAD2.<sup>75</sup> MAD2, in turn, is known to undergo  $\beta$ -sheet augmentation with reshuffling upon complex formation with CDC20 (cell-division-control protein 20) or MAD1.<sup>110</sup> This interaction motif, however, brings the risk of aggregation. In  $\beta$ 2-microglobulin, the pathogenic D76N mutation destabilizes protective edge  $\beta$ -strands, which exposes aggregation-prone regions and enables amyloid formation.<sup>111</sup>

In biomolecular MAS NMR, the observation of a residue is indicative of slow dynamics associated with a regular, stable structure. In the case of the complex of N-FAT10-C0 and NUB1L, this applies to residues I68-V81, of which T73-V81 form a  $\beta$ -strand, the N-terminus (G4), and eleven more residues of which we could sequentially assign only W17 (there is just one Trp in the sequence). The other, unobserved residues of N-FAT10-C0 are disordered and quite possibly dynamic on micro- to millisecond time scales – motion in this intermediate regime tends to make both J-coupling and dipolar-coupling based experiments inefficient. Integrated approaches of MAS NMR, liquid-state NMR, and molecular dynamics simulations, have demonstrated such dynamic disorder in enzymes, its connection to structural frustration, and the functional importance.<sup>112,113</sup> Based on the MAS NMR data and the AlphaFold-Multimer prediction (see below) we propose that NUB1L and N-FAT10 form a fuzzy complex<sup>114</sup> in which NUB1L is the predominantly folded “reverse chaperone” and N-FAT10 is the predominantly unfolded client. The two proteins interact *via* the intermolecular  $\beta$ -sheet and *via* electrostatic and hydrophobic interactions at the scattered, non-sequentially assigned residues. The N-terminus of N-FAT10 binds to NUB1L in a yet-unknown location.

AlphaFold-Multimer predicts with high confidence that the  $\beta$ 2-strand and  $\alpha$ 1-helix of N-FAT10-C0 remain partially intact (Figures 2.5c, S2.11), but this is not reflected in the MAS NMR spectra. This likely relates to the known tendency of AlphaFold to predict, for intrinsically disordered proteins or regions, the structures of conditionally folded states, i.e., stable structures that form under specific conditions such as the interaction with a binding partner or following post-translational modification.<sup>115</sup> For example, AlphaFold’s predicted structure of isolated 4E-BP2 simultaneously contains the four  $\beta$ -strand

structure that forms upon multisite phosphorylation<sup>116</sup> and two helices that resemble those observed in crystal structures of 4E-BP2 and 4E-BP1 bound to translation initiator factor 4E.<sup>117,118</sup> The prediction of the intact  $\beta$ 2-strand and  $\alpha$ 1-helix is thus likely a remnant of the conditional folding of isolated N-FAT10-C0. There is additional disagreement between AlphaFold-Multimer and the MAS NMR data regarding the length of the intermolecular  $\beta$ -sheet. AlphaFold-Multimer predicts residues L76-V81 of N-FAT10-C0 to participate whereas TALOS-N classifies T73-V81 as  $\beta$ -strand. This is not just due to the rendering of the secondary structure: the backbone torsion angles of I74 are clearly distinct (TALOS-N:  $\Phi = -133.3^\circ$ ,  $\Psi = 155.3^\circ$ ; AlphaFold-Multimer/VADAR:  $\Phi = -117.1^\circ$ ,  $\Psi = -24.2^\circ$ , see Figure S2.10). To accommodate the longer  $\beta$ -sheet, NUB1L would have to undergo a conformational change. Finally, AlphaFold-Multimer does not predict the interactions of NUB1L with the scattered anchor residues of N-FAT10-C0, including the N-terminus. All this emphasizes that experimental data remain essential in the investigation of the structures and functions of weakly folded proteins. MAS NMR experiments that could shed more light on the interaction of N-FAT10 and NUB1L include frequency-selective rotational-echo double-resonance (REDOR)<sup>29</sup> to measure specific  $^{13}\text{C}$ - $^{15}\text{N}$  distances, the double-REDOR filter<sup>119</sup> to identify intermolecular interfaces,  $^1\text{H}$  detection enabled by ultrafast MAS<sup>39</sup> to explore the hydrogen bonding network of the anti-parallel  $\beta$ -sheet, and relaxation dispersion<sup>120</sup> to probe microsecond time-scale conformational fluctuations.

## 2.6 Conclusion

We have reported the MAS NMR investigation of the ubiquitin-like modifier FAT10 and its interaction with adapter protein NUB1L. Sequential assignment of amino acid residues followed by empirical prediction of torsion angles and secondary structure has provided insight into the biological function of the weak folding of FAT10. When the N-domain of FAT10 is in the ubiquitin-like  $\beta$ -grasp fold, residues I68 through (approximately) V80 constitute a partially unstructured degron. In preparation of rapid degradation by the proteasome, NUB1L stabilizes the N-domain of FAT10 in a mostly unfolded state. Residues I68-V81 attain structural homogeneity and so do a number of anchor residues, including the N-terminus. Chemical shifts indicate that residues T73-V81 form a  $\beta$ -strand, providing experimental, structural evidence for  $\beta$ -strand capture in ubiquitin-independent protein degradation. The binding of the N-terminus, presumably to NUB1L, likely plays a critical role in the presentation of the unfolded N-FAT10 chain to the proteasome.

## 2.7 Supporting Information

### 2.7.1 Protein expression and purification

#### **$\text{U-}^{13}\text{C}$ , $^{15}\text{N}$ -N-FAT10-C0**

Cysteine-free N-domain of human FAT10 (amino acids 5-86; C7T, C9T) was expressed as a His<sub>6</sub>-GST-fusion protein in *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells (Agilent Technologies). The pETM-30 plasmid was kindly provided by the Institute of Cell Biology and Immunology Thurgau (Kreuzlingen, Switzerland); a TEV protease cleavage site C-terminal to the purification tag results in an

additional glycine residue at the N-terminus after cleavage. For uniform  $^{13}\text{C}$  and  $^{15}\text{N}$  labelling, 3.6 g of  $\text{U-}^{13}\text{C}_6\text{-D-glucose}$  and 0.5 g of  $^{15}\text{N}$ -ammonium chloride per liter of M9 minimal medium were added as exclusive sources of carbon and nitrogen. Bacteria cells were grown at 37 °C to an  $\text{OD}_{600}$  of 0.5-0.6, induced with 0.4 mM IPTG at 21 °C overnight, and harvested by centrifugation (RCF 4000 g, 10 min, 8 °C). Harvested cells were lysed in lysis buffer (20 mM TRIS-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole, 10 % v/v glycerol, 0.1 % v/v Triton X-100, 100  $\mu\text{g}/\text{mL}$  lysozyme, 1 mM PMSF and EDTA-free protease inhibitor). After sonication, cell debris was removed by centrifugation (RCF 47000 g, 30 min, 8 °C). The supernatant was filtered and applied to Ni-NTA Agarose beads (Macherey-Nagel) for affinity chromatography. The protein was eluted with elution buffer (20 mM TRIS-HCl (pH 8.0), 300 mM NaCl and 500 mM imidazole) and buffer exchanged to binding buffer (20 mM TRIS-HCl (pH 8.0), 300 mM NaCl and 10 mM imidazole). His-tagged TEV protease was allowed to act overnight (5  $\mu\text{g}$  of protease per 1 mg of recombinant protein). After cleavage, His-tagged TEV protease and cleavage byproducts were separated by a second Ni-affinity chromatography using binding buffer. The volume of the flow-through containing the  $\text{U-}^{13}\text{C},^{15}\text{N}$ -FAT10-C0 was reduced to 5 mL and filtered. Further purification was achieved by size-exclusion chromatography (Cytiva, formerly Amersham Biosciences, ÄKTA pure chromatography system equipped with a HiLoad 16/600 Superdex 75 pg column) yielding up to 7 mg per liter of M9 medium. Before preparation of the samples for MAS NMR,  $\text{U-}^{13}\text{C},^{15}\text{N}$ -FAT10-C0 was in 20 mM HEPES (pH 7.5) and 150 mM NaCl and concentrated to specified values (Amicon Ultra-15 centrifugal filter unit, MWCO 3 kDa).

### **NUB1L**

Human NUB1L (amino acids 2-615) was expressed as a His<sub>6</sub>-SUMO-fusion protein also in *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells. The pSUMO plasmid was kindly provided by the Institute of Cell Biology and Immunology Thurgau (Kreuzlingen, Switzerland). Bacteria cells were grown in LB medium at 37 °C to an  $\text{OD}_{600}$  of 0.7, induced with 0.4 mM IPTG at 21 °C overnight and harvested by centrifugation (RCF 4000 g, 10 min, 8 °C). The harvested cells were lysed in lysis buffer (20 mM TRIS-HCl (pH 7.5), 150 mM NaCl, 20 mM imidazole, 1 mM TCEP, 10 % v/v glycerol, 0.1 % v/v Triton X-100, 100  $\mu\text{g}/\text{mL}$  lysozyme, 1 mM PMSF and EDTA-free protease inhibitor). After sonication, cell debris was removed by centrifugation (RCF 27000 g, 30 min, 8 °C). The supernatant was filtered and applied to Ni-NTA Agarose beads for affinity chromatography. The protein was eluted with elution buffer (20 mM TRIS-HCl (pH 7.5), 150 mM NaCl, 500 mM imidazole and 1 mM TCEP) and buffer exchanged to binding buffer (20 mM TRIS-HCl (pH 7.5), 150 mM NaCl, 20 mM imidazole and 1 mM TCEP). His-tagged Ulp1 protease was added overnight (4  $\mu\text{g}$  of protease per 1 mg of recombinant protein). After cleavage, His-tagged Ulp1 protease and cleavage byproducts were separated by a second Ni-affinity chromatography using binding buffer. The volume of the flow-through containing (natural abundance) NUB1L was reduced to 5 mL and filtered. Further purification was achieved by size-exclusion chromatography yielding up to 15 mg of pure NUB1L per liter of LB medium. To form the complex with N-FAT10-C0, the protein was in 20 mM HEPES (pH 7.5), 150 mM NaCl, and 1 mM TCEP, concentrated to approximately 20 mg/mL (Amicon Ultra-15 centrifugal filter unit, MWCO 30 kDa).

## 2.7.2 Preparation of samples for MAS NMR

### **Microcrystalline N-FAT10-C0**

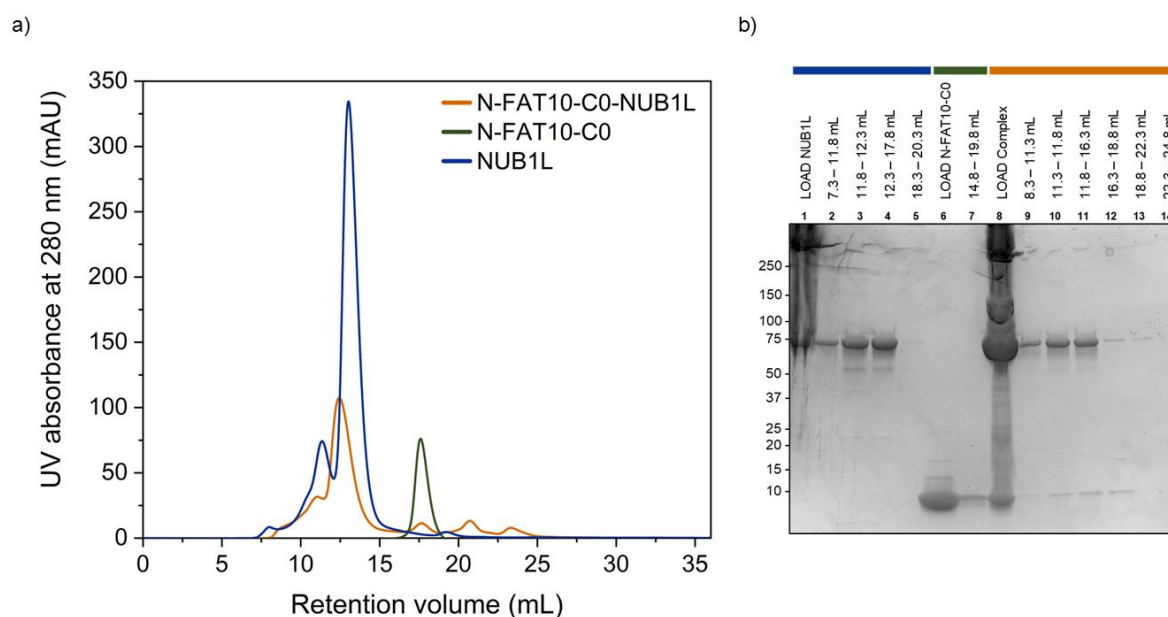
Microcrystals of U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -N-FAT10-C0 were grown with the sitting drop vapour diffusion method. Equal volumes of protein at a concentration of approximately 20 mg/mL and crystallization solution (3.2 M  $(\text{NH}_4)_2\text{SO}_4$  and 0.1 M citric acid (pH 2.5)) were mixed in drops of 25  $\mu\text{L}$ . Each drop was equilibrated against 0.75 mL of crystallization solution supplemented with additional 150 mM NaCl. After keeping the crystallization trays at 4  $^\circ\text{C}$  for four days, microcrystals were harvested. Ultracentrifugation was used to assure dense packing in a 3.2 mm MAS rotor. A home-made packing tool containing the protein suspension was centrifuged for 2 h at RCF 210000 g and 4  $^\circ\text{C}$ . The supernatant was removed and packing efficiency was assessed by UV-Vis absorbance measurements of the supernatant at 280 nm. The rotor, filled with an estimated 13.5 mg of microcrystalline U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -N-FAT10-C0, was closed with a Vespel cap.

### **Lyophilized/rehydrated N-FAT10-C0**

After concentration to approximately 8 mg/mL, U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -N-FAT10-C0 was lyophilized, ground, and packed into a 3.2 mm MAS rotor. The rotor containing an estimated 14.5 mg of protein was closed with a Vespel cap. Rehydration was accomplished by adding Milli-Q water in amounts of 1  $\mu\text{L}$  while monitoring the  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization spectrum. After addition of a total of 7  $\mu\text{L}$ , no further improvement in spectral resolution was observed.

### **N-FAT10-C0 in complex with NUB1L**

Complex formation of N-FAT10-C0 and NUB1L was verified with size-exclusion chromatography (Figure S2.1). To prepare the complex for MAS NMR spectroscopy, co-sedimentation into the rotor was used. 1 mL of U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -N-FAT10-C0 at a concentration of 3 mg/mL was mixed with 1 mL of natural abundance NUB1L with a concentration of 20 mg/mL, leading to a molar ratio of 1:1. The packing tool containing this solution was centrifuged at RCF 210000 g and 4  $^\circ\text{C}$ . As the holding capacity of the packing tool is only 1 mL, the run was interrupted to remove supernatant and add more of the original solution over the course of several days. Packing efficiency was assessed by UV-Vis absorbance measurements of the supernatant at 280 nm. The 3.2 mm rotor was closed with a Vespel cap.



**Figure S2.1:** Size-exclusion chromatography, on a Cytiva ÄKTA pure chromatography system equipped with a Superdex 200 Increase 10/300 GL column, and subsequent SDS-PAGE analysis. (a) Individual runs of isolated N-FAT10-C0 (green) and isolated NUB1L (dark blue) are shown together with a mixture of both N-FAT10-C0 and NUB1L (orange). Molar amounts were adjusted to be the same in each run. The shift to smaller retention volumes in the orange chromatogram indicates a larger molecular weight and thereby complex formation. (b) SDS-PAGE analysis before and after size-exclusion chromatography. 5  $\mu$ L of gel sample buffer with SDS (reducing conditions) was added to 20  $\mu$ L of each sample (loads for size-exclusion chromatography as well as collected fractions containing protein). After boiling for 5 min at 95  $^{\circ}$ C, samples were loaded on a gradient gel and separated in 30 min by applying a constant voltage of 200 V, followed by colloidal Coomassie staining. For molecular weight estimation, Precision Plus Protein Dual Color Standards (Bio-Rad) were used.

### 2.7.3 MAS NMR spectroscopy

MAS NMR experiments were performed at 18.8 T ( $^1\text{H}$  Larmor frequency of 800.3 MHz) in the NMR Core Facility at the University of Konstanz and at 20.0 T ( $^1\text{H}$  Larmor frequency of 850.2 MHz) at ETH Zürich. The console was Bruker Avance NEO or Avance III. Bruker 3.2 mm E-free HCN MAS probes were used, together with regular-wall 3.2 mm ZrO rotors (maximum sample volume 32.1  $\mu$ L). During experiments, the sample was kept at a temperature of approximately 4  $^{\circ}$ C, unless noted otherwise. Calibration of the Bruker Cooling Unit (BCU) II was performed by monitoring the chemical shift of  $^{79}\text{Br}$  in KBr powder.<sup>121</sup> The carbon dimension was referenced indirectly to DSS in  $\text{D}_2\text{O}$  (0.5 % by weight),<sup>122</sup> i.e., the  $^{13}\text{C}$  adamantane methylene peak is observed at 40.49 ppm. The nitrogen dimension was referenced to liquid ammonia at 25  $^{\circ}$ C.<sup>123</sup> Table S2.1 provides an overview of the experiments performed per sample.

**Table S2.1:** MAS NMR spectra obtained for each sample.

Experiment	Field (T)	Number of scans	$^{15}\text{N}$ - $^{13}\text{C}$ / $^{13}\text{C}$ - $^{13}\text{C}$ mixing (ms)	Temperature ( $^{\circ}\text{C}$ )	MAS frequency (kHz)	Measurement time
<b>Microcrystalline U-<math>^{13}\text{C}</math>, <math>^{15}\text{N}</math>-N-FAT10-C0</b>						
hC	18.8	256		4	14.5	
hN	18.8	128		4	14.5	
DARR	18.8	32	10	4	14.5	16 h
ZF TEDOR	18.8	32	0.52	4	14.5	7 h
ZF TEDOR-DARR	18.8	4	1.1/40	4	14.5	7 d 21 h
<b>Lyophilized/rehydrated U-<math>^{13}\text{C}</math>, <math>^{15}\text{N}</math>-N-FAT10-C0</b>						
hC	18.8	256		4	14.5	
hN	18.8	512		4	19.0	
DARR	18.8	32	10	4	14.5	22 h
ZF TEDOR	18.8	128	0.84	4	19.0	15 h
ZF TEDOR-DARR	18.8	8	0.84/40	4	19.0	5 d 22 h
NCOCX	20.0	32 (2x)	55.6 (CORD)	4	19.0	3 d (2x)
<b>U-<math>^{13}\text{C}</math>, <math>^{15}\text{N}</math>-N-FAT10-C0 in complex with natural abundance NUB1L</b>						
hC	20.0	128/128/128, 256		-10/-2/4	19.0	
hN	20.0	512		4	19.0	
DARR	20.0	256	10	4	19.0	3 d
ZF TEDOR	20.0	256 (2x)	1.26	4	19.0	1 d 12 h (2x)
ZF TEDOR-DARR	20.0	8 (3x)	1.26/50	4	19.0	5 d 5 h (3x)

Typical  $\pi/2$ -pulse lengths were 3.0-3.5  $\mu\text{s}$  for  $^1\text{H}$ , 3.0-5.0  $\mu\text{s}$  for  $^{13}\text{C}$  and 6.0-10.5  $\mu\text{s}$  for  $^{15}\text{N}$ . The  $^1\text{H}$ - $^{13}\text{C}/^1\text{H}$ - $^{15}\text{N}$  cross-polarization (CP)<sup>124</sup> step was realized using a linearly ramped RF field on the  $^1\text{H}$  channel (either from 70 or 90 to 100 % amplitude); the centre of the ramp was set to match the  $n = +1$  Hartmann-Hahn condition.<sup>20,125</sup> Swept-frequency two-pulse phase modulation (SW<sub>r</sub>-TPPM)<sup>24</sup> or small phase incremental alternation with 64 steps (SPINAL-64)<sup>25</sup> were applied for proton decoupling. The States-TPPI method<sup>126</sup> was used for phase-sensitive detection in the indirect dimensions.

Before running a  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  ZF TEDOR-DARR experiment, a  $^{15}\text{N}$ - $^{13}\text{C}$  ZF TEDOR experiment served to set the acquisition parameters correctly, including the mixing time and the duration of the REDOR  $\pi$ -pulses. The z-filter delays were set to  $\sim 200$   $\mu\text{s}$ , matching a multiple of the rotor period. Table S2.2 gives a complete overview of the acquisition parameters per experiment and sample.

Chapter 2 Intermolecular  $\beta$ -sheet formation guides interaction between FAT10 and NUB1L

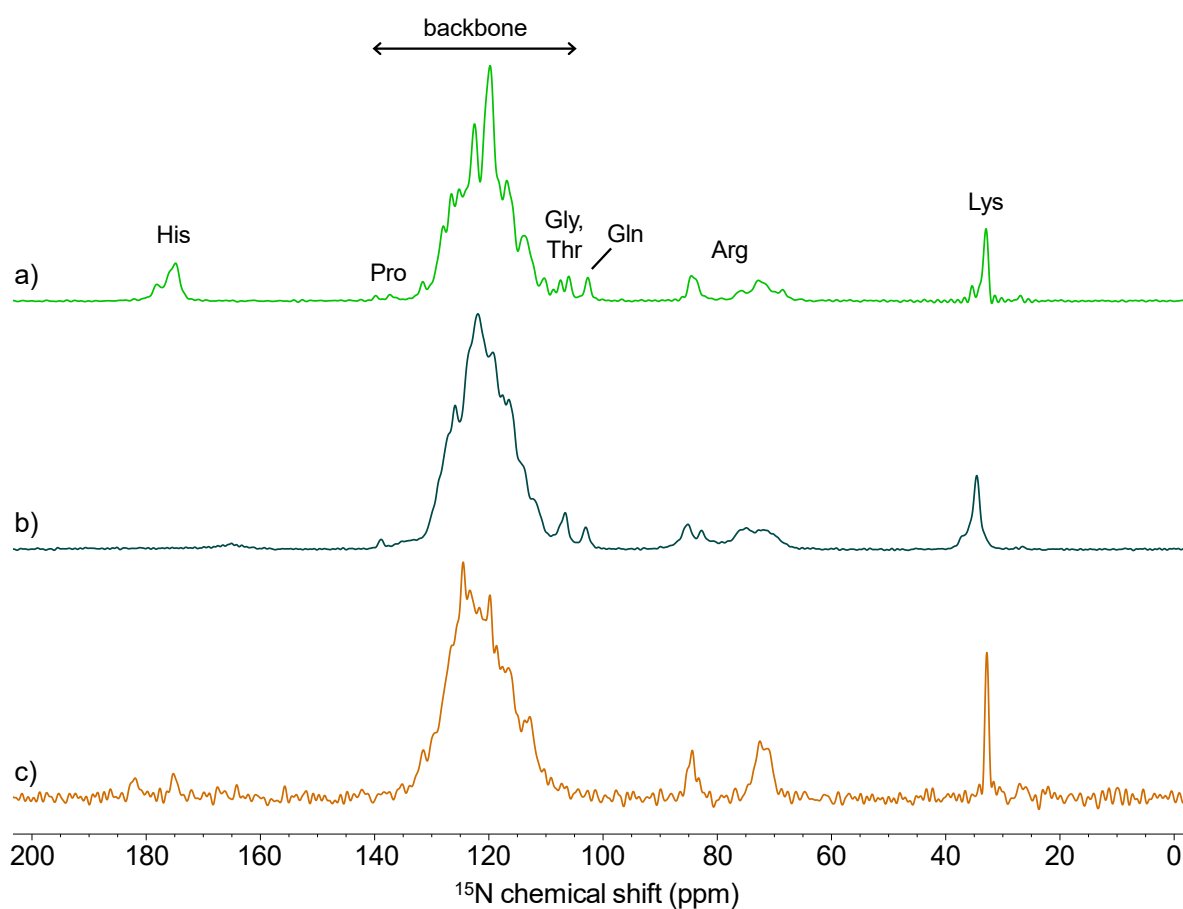
**Table S2.2:** Acquisition parameters for MAS NMR experiments.

Experiment	CP contact time (ms)		Radio frequency field strength (kHz)							Dwell time ( $\mu$ s)			Acquisition time (ms)			Carrier position (ppm)			Recycle delay (s)	
	$^1\text{H}$ - $^{13}\text{C}$ / $^1\text{H}$ - $^{15}\text{N}$ CP	$^{15}\text{N}$ - $^{13}\text{C}$ DCP	$^1\text{H}$ - $^{13}\text{C}$ / $^1\text{H}$ - $^{15}\text{N}$ CP		$^{15}\text{N}$ - $^{13}\text{C}$ DCP		$\pi/2$ - and $\pi$ -pulses		Decoupling											
			$^1\text{H}$ (linear) <sup>#</sup>	$^{13}\text{C}/^{15}\text{N}$	$^{15}\text{N}$	$^{13}\text{C}$ (tangent) <sup>#</sup>	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$	$^1\text{H}$	$^{13}\text{C}$	$^{13}\text{C}$	$^{15}\text{N}$	$^{13}\text{C}$	$^{13}\text{C}$	$^{15}\text{N}$	$^{13}\text{C}$	$^{13}\text{C}$		$^{15}\text{N}$
<b>Microcrystalline U-<math>^{13}\text{C}</math>, <math>^{15}\text{N}</math>-N-FAT10-C0</b>																				
hC	0.9		67.9	50.0			83.3			71.4, SW <sub>r</sub> -TPPM	8.5			17.4			100			4.0
hN	1.0		74.8	50.0			71.4			71.4, SPINAL-64			12.0			12.3			104	4.0
DARR	1.0		75.9	50.0			83.3	50.0		71.4, SPINAL-64	11.3	23.0		23.1	6.9		100	100		3.0
ZF TEDOR	0.9		67.9	50.0			83.3	50.0	23.8	71.4, SW <sub>r</sub> -TPPM	12.0		69.0	13.2		10.3	100		104	2.5
ZF TEDOR-DARR	0.9		67.9	50.0			83.3	50.0	23.8	71.4, SW <sub>r</sub> -TPPM/ SPINAL-64	11.3	23.0	69.0	13.6	2.9	8.8	100	100	104	2.5
<b>Lyophilized/rehydrated U-<math>^{13}\text{C}</math>, <math>^{15}\text{N}</math>-N-FAT10-C0</b>																				
hC	0.9		67.9	50.0			83.3			71.4, SPINAL-64	6.1			25.0			100			3.0
hN	0.7		64.3	41.7			83.3			71.4, SPINAL-64			10.0			20.5			104	3.0
DARR	0.9		67.9	50.0			83.3	50.0		71.4, SPINAL-64	11.3	23.0		23.1	9.3		100	100		3.0
ZF TEDOR	0.8		72.2	50.0			83.3	83.3	36.2	90.6/83.3, SW <sub>r</sub> - TPPM	12.0		52.6	24.6		5.3	100		87	2.0
ZF TEDOR-DARR	0.9		72.2	50.0			83.3	83.3	36.2	90.6/83.3, SW <sub>r</sub> - TPPM; 83.3, SPINAL-64	6.1	52.6	105.3	12.5	5.0	8.4	50	50	87	2.0
NCOCX	0.7	1.3	79.5 (tangent)	41.7	24.0	56.2	83.3	62.5	41.7	90.0, SW <sub>r</sub> -TPPM; 100 kHz, CW	11.0	210.5	210.5	22.5	5.1	8.4	85	85 <sup>†</sup>	122	2.0
<b>U-<math>^{13}\text{C}</math>, <math>^{15}\text{N}</math>-N-FAT10-C0 in complex with natural abundance NUB1L</b>																				
hC	1.0		72.7	50.0			83.3			83.3, SPINAL-64/ SW <sub>r</sub> -TPPM*	5.8			11.9			100			3.0
hN	0.8		64.5	41.7			83.3			96.2, SW <sub>r</sub> -TPPM			11.0			22.5			104	3.0
DARR	1.0		72.7	50.0			83.3	50.0		83.3, SPINAL-64	11.0	21.1		11.3	5.2		100	100		2.0
ZF TEDOR	1.0		72.7	50.0			83.3	62.5	38.5	96.2/83.3, SW <sub>r</sub> - TPPM	11.0		52.6	11.3		6.5	100		104	2.0
ZF TEDOR-DARR	1.0		72.7	50.0			83.3	62.5	38.5	96.2/83.3, SW <sub>r</sub> - TPPM; 83.3, SPINAL-64	6.1	52.6	105.3	12.5	5.0	7.4	56	56	83	2.0

<sup>#</sup>Corresponds to 100 % of the ramp. \*Decoupling sequence for  $^1\text{H}$ - $^{13}\text{C}$  CP temperature series. <sup>†</sup>For  $^{15}\text{N}$ - $^{13}\text{C}$  SPECIFIC-CP, the CO carrier was centred at 174 ppm.

For free induction decays with unnecessary long acquisition times, the number of effective points was adjusted accordingly and zero-filling was applied. 1D spectra were baseline corrected, but no apodization was used. For 2D and 3D spectra, 45°- or 72°-shifted squared-sine bell apodization (corresponding to an SSB value of 4 or 2.5 in TopSpin and an offset of 0.25 or 0.4 in NMRPipe) was used. The resolution of the ZF TEDOR-DARR spectrum of microcrystalline N-FAT10-C0 was improved by forward linear prediction to 1.5 times the number of the original data points in the indirect carbon dimension.

#### 2.7.4 $^1\text{H}$ - $^{15}\text{N}$ cross-polarization spectra



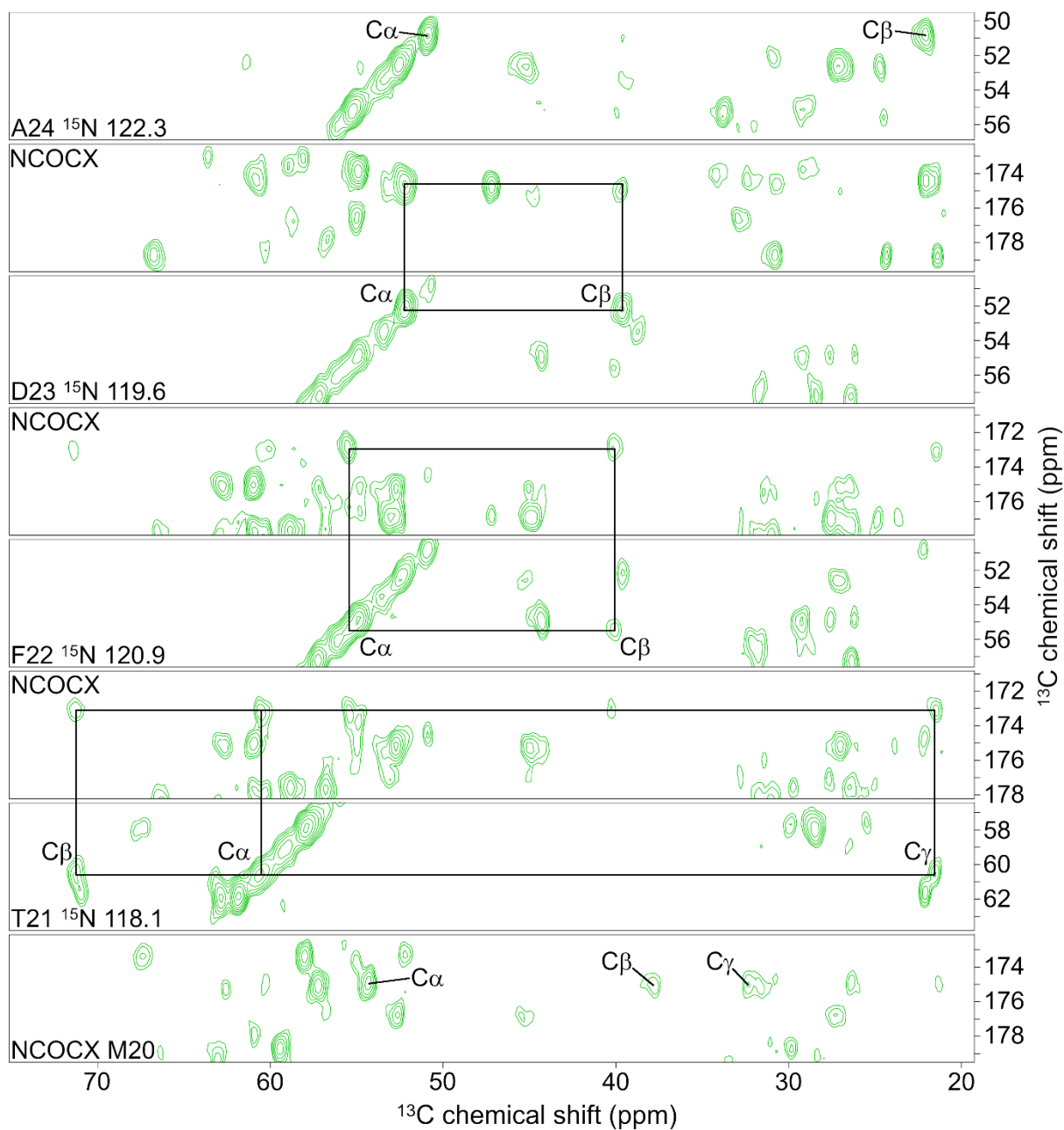
**Figure S2.2:**  $^1\text{H}$ - $^{15}\text{N}$  cross-polarization spectra of (a) microcrystalline N-FAT10-C0, (b) lyophilized/rehydrated N-FAT10-C0, and (c) N-FAT10-C0 in complex with NUB1L.

2.7.5  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  spectra of microcrystalline N-FAT10-C0 and visualization of secondary structure predictions

**Figure S2.3 (on the following page):** (a)  $^{13}\text{C}$ - $^{13}\text{C}$  DARR spectrum of microcrystalline N-FAT10-C0. Sequentially assigned cross peaks are marked. To maintain readability, only labels are included of residues that are also observed in the  $^{13}\text{C}$ - $^{13}\text{C}$  DARR spectrum of lyophilized/rehydrated N-FAT10-C0, i.e., cross peaks from D18, F22, D23, H36, V45, Q46, K55, K58, R60, R61, T73, H75, L76, K79, V80 and V81 are not labelled. (b)  $^{15}\text{N}$ - $^{13}\text{C}$  ZF TEDOR spectrum of microcrystalline N-FAT10-C0. Sequentially assigned cross peaks are marked and labelled. Cross peaks from side chains of His ( $^{15}\text{N}$  chemical shift 170-180 ppm), Arg ( $^{15}\text{N}$  chemical shift 65-85 ppm) and Lys ( $^{15}\text{N}$  chemical shift 30-35 ppm) are observed, but not plotted here. (c) Projection of predicted secondary structure per residue on the X-ray structure of N-FAT10-C0 (PDB-ID 6GF1, chain B). H stands for helix, E for strand, and L for loop; predictions with a confidence > 0.75 are highly reliable. Residues K32 and K34 are ambiguously assigned, but secondary structure is predicted based on observed chemical shifts. For G67, secondary structure is predicted based on the sequence, but this residue is color coded as the assignment is unambiguous. D69 and T73 are ambiguously assigned, but colored yellow as secondary structure prediction is based on the sequence. Residue I74 is unassigned, but the secondary structure is still predicted based on chemical shifts of neighboring residues.

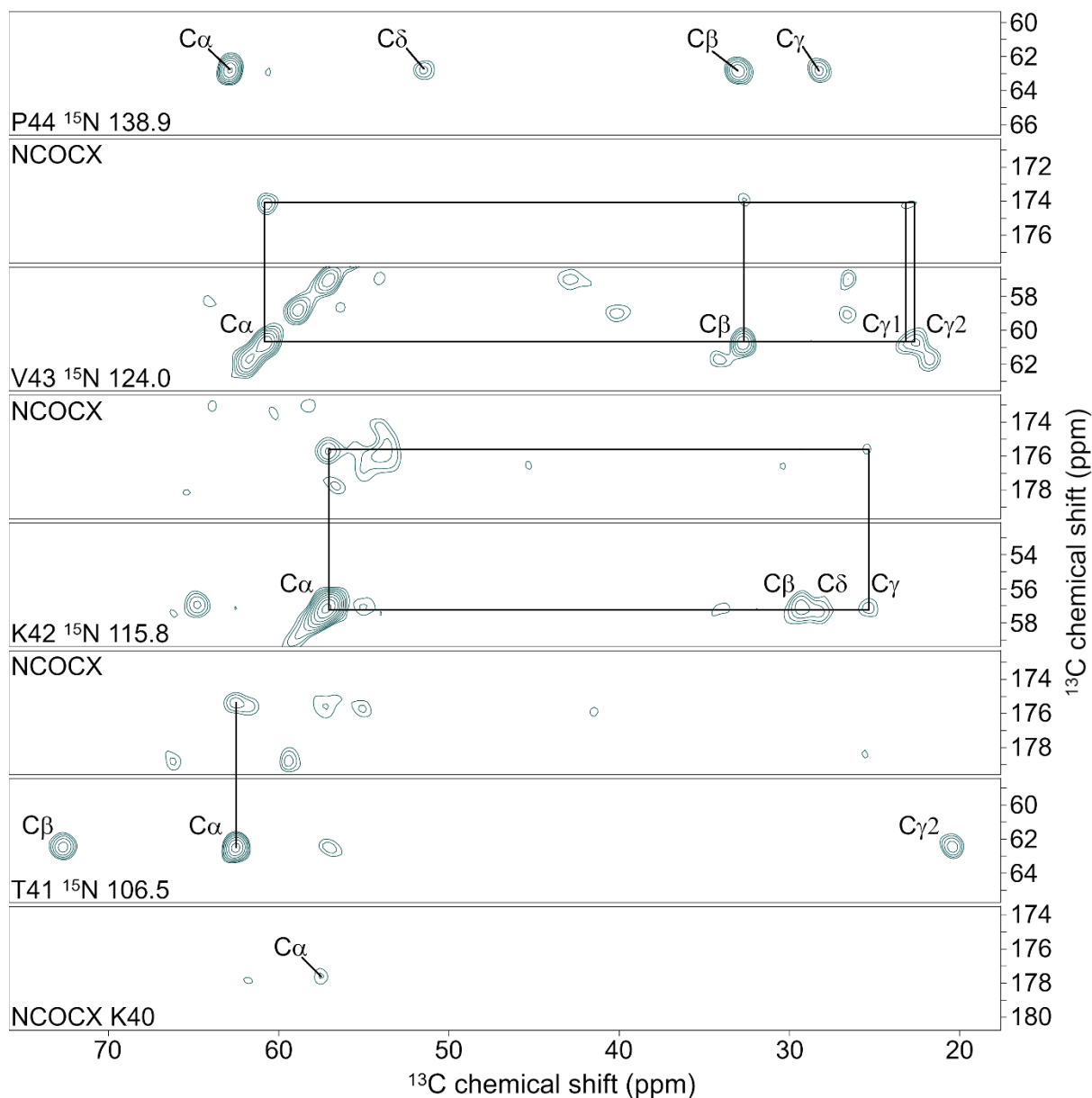


## 2.7.6 Strip plot of microcrystalline N-FAT10-C0

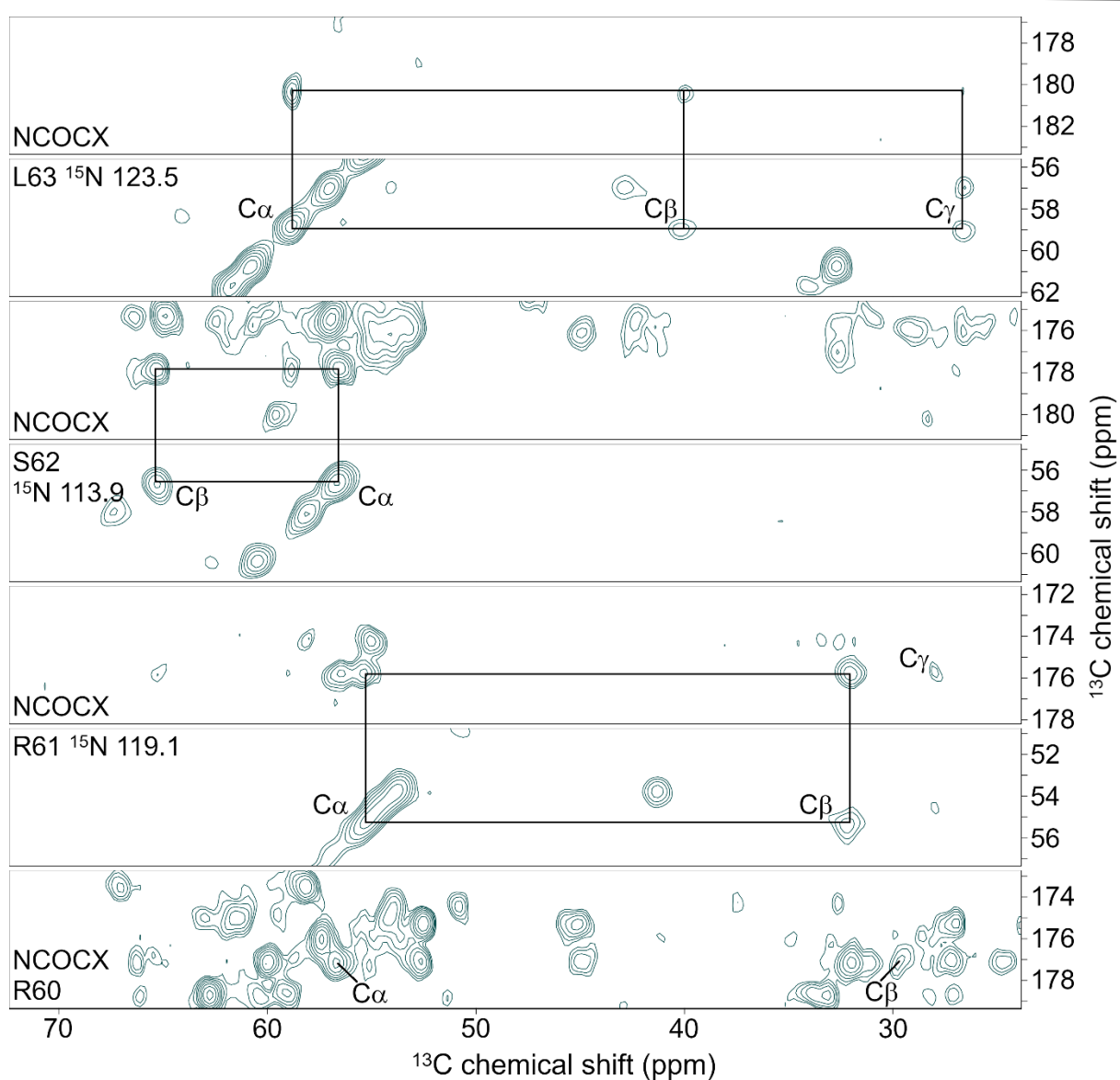


**Figure S2.4:** Representative strip plot from the  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  ZF TEDOR-DARR spectrum of microcrystalline N-FAT10-C0. The 120-126 ppm  $^{15}\text{N}$  planes are particularly crowded.

## 2.7.7 Strip plots of lyophilized/rehydrated N-FAT10-C0

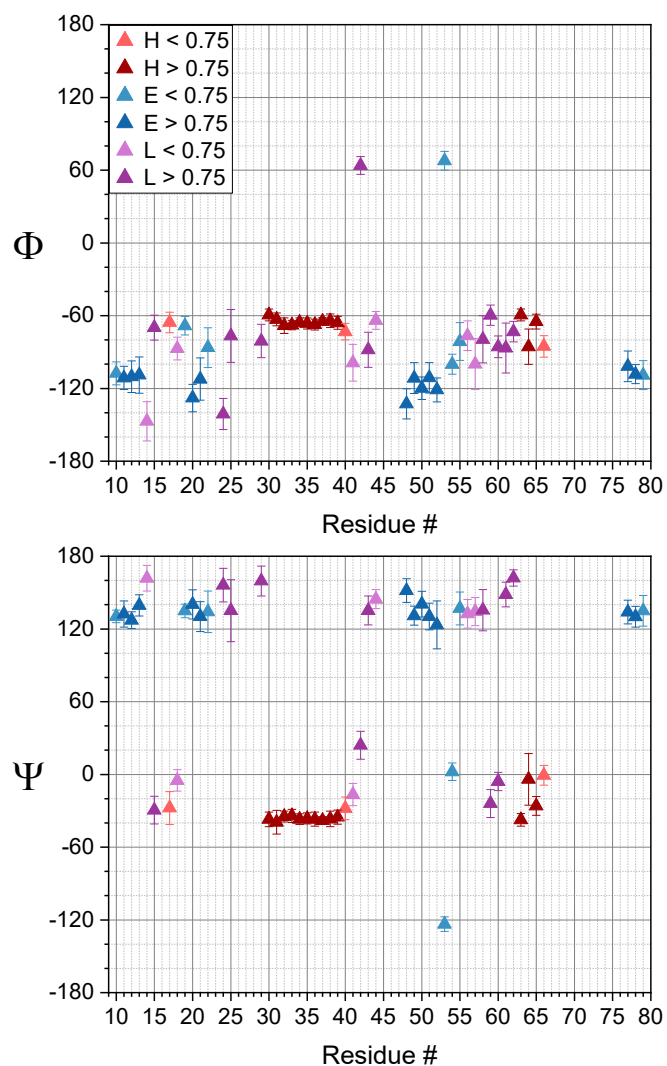


**Figure S2.5:** Representative strip plot from the  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  ZF TEDOR-DARR spectrum of lyophilized/rehydrated N-FAT10-C0.



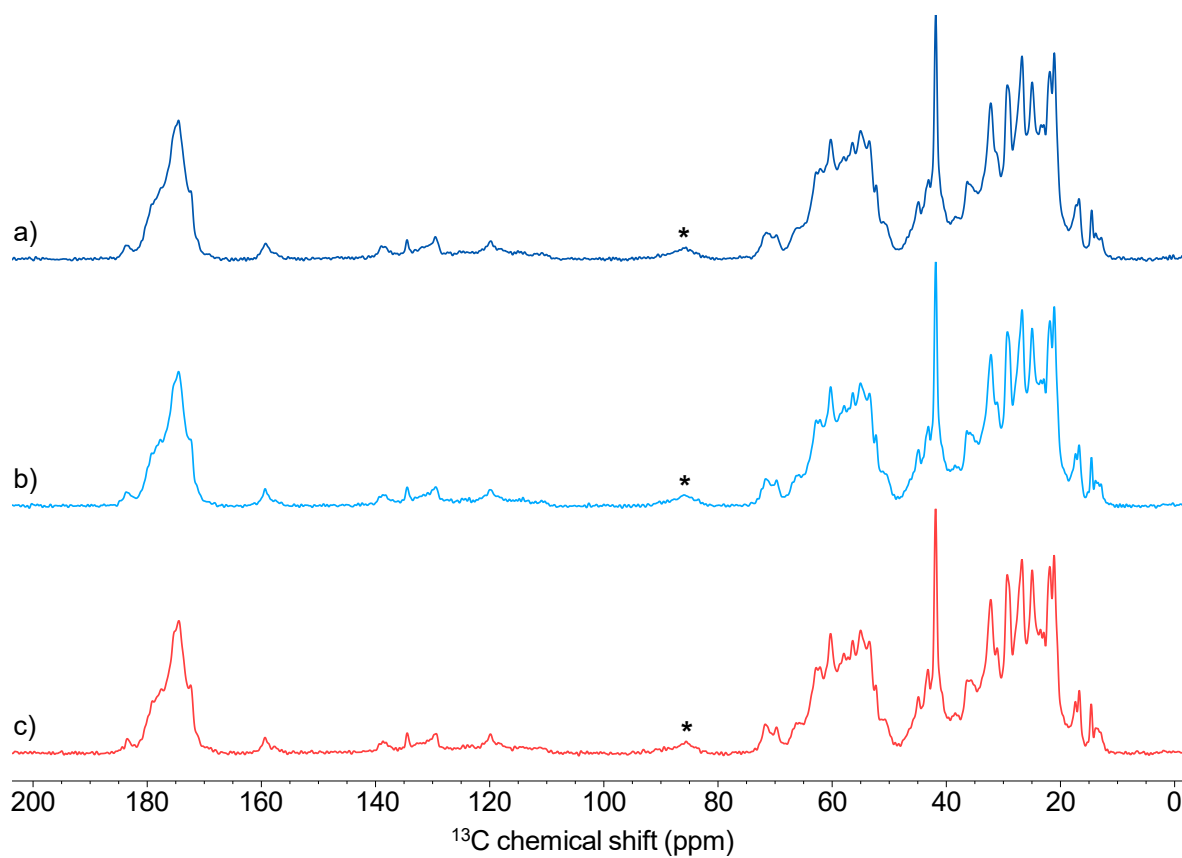
**Figure S2.6:** Representative strip plot from the  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  SPECIFIC-CP/CORD spectrum of lyophilized/rehydrated N-FAT10-C0.

## 2.7.8 Torsion angle analysis of lyophilized/rehydrated N-FAT10-C0



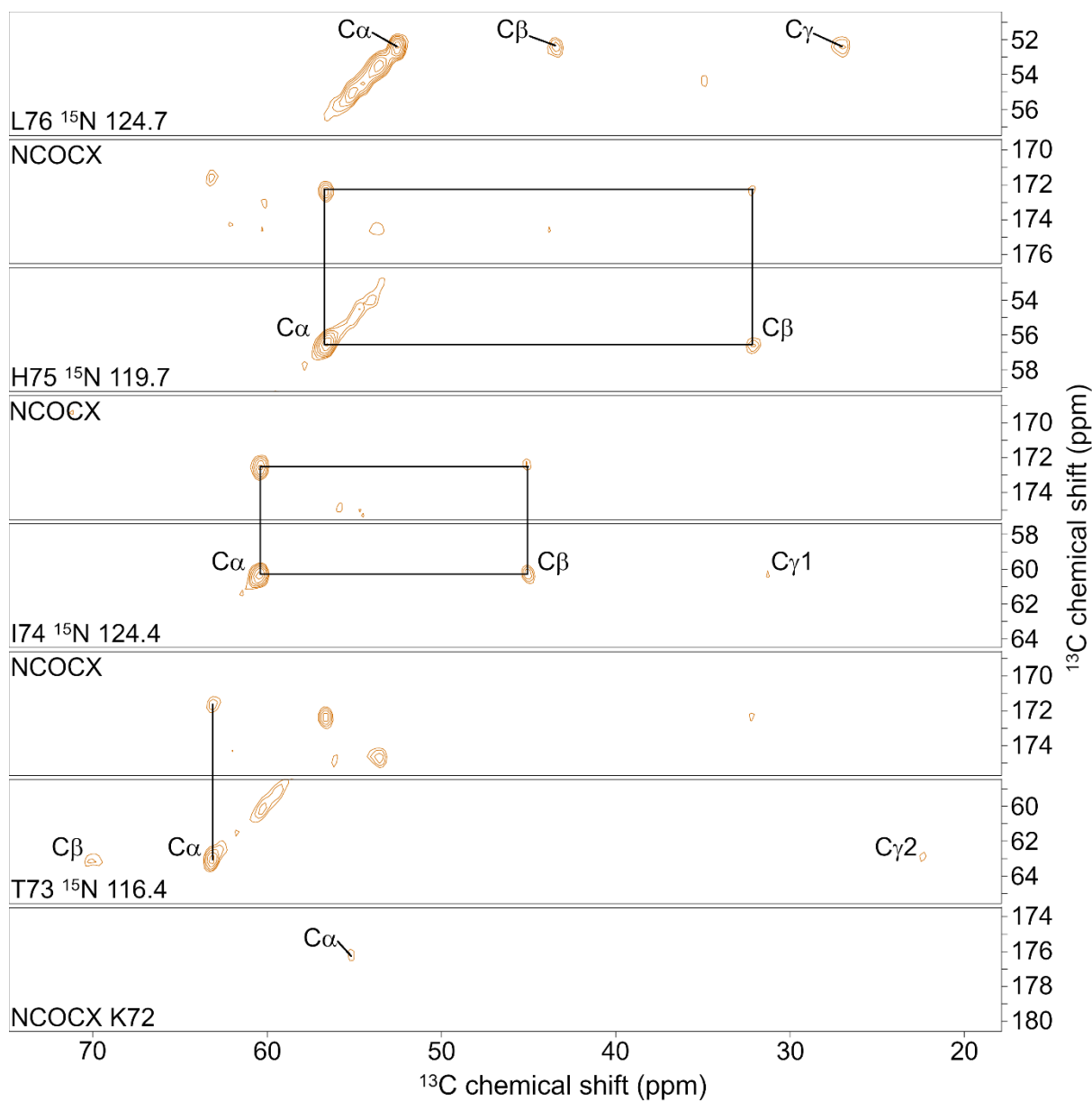
**Figure S2.7:** Backbone torsion angle and secondary structure prediction based on MAS NMR of lyophilized/rehydrated N-FAT10-C0. Error bars correspond to the standard deviations of the  $\Phi$  and  $\Psi$  angles of the best matches in the TALOS-N database. For residues E16, D23, Y27, D28, D47, and G67 there is no consensus among the database matches and, hence, no predicted torsion angles are shown. No predictions are made for T9, P26, V45, Q46, I68-L76, and V80. Secondary structure predictions are based on observed chemical shifts and classified as helix (H), strand (E), or loop (L). Predictions with a confidence > 0.75 are highly reliable.

2.7.9 Temperature dependence of N-FAT10-C0 in complex with NUB1L



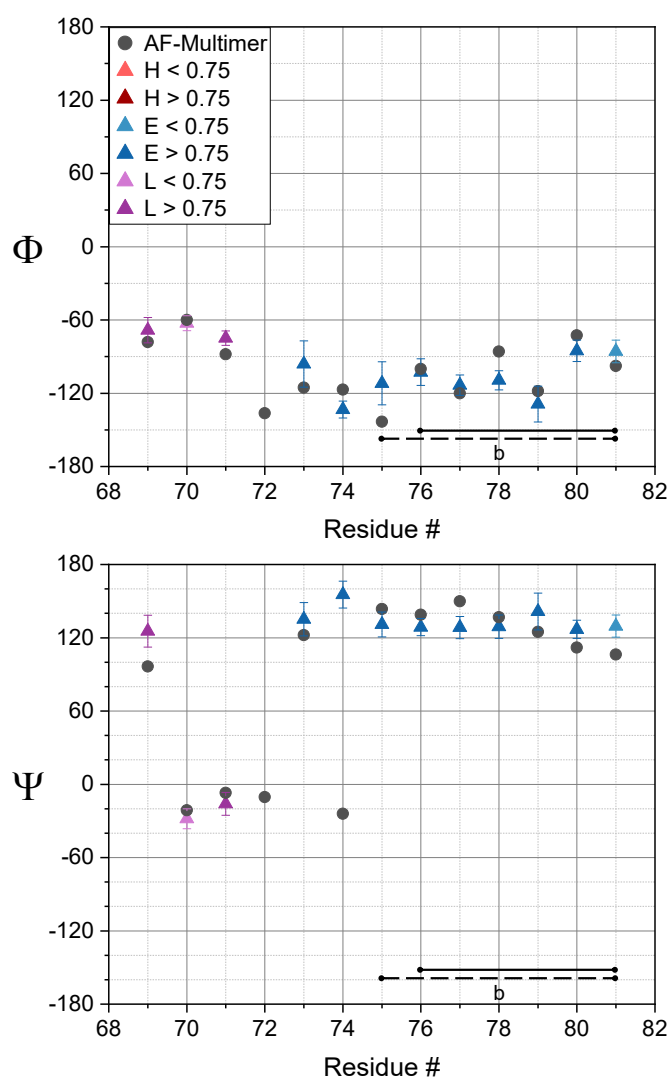
**Figure S2.8:**  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization spectra of N-FAT10-C0 in complex with NUB1L at temperature of (a) -10 °C, (b) -2 °C, and (c) 4 °C.

## 2.7.10 Strip plot of N-FAT10-C0 in complex with NUB1L



**Figure S2.9:** Representative strip plot from the  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  ZF TEDOR-DARR spectrum of N-FAT10-C0 co-sedimented with NUB1L.

## 2.7.11 Torsion angle analysis of N-FAT10-C0 in complex with NUB1L



**Figure S2.10:** Backbone torsion angle and secondary structure prediction based on MAS NMR of N-FAT10-C0 co-sedimented with NUB1L. Error bars correspond to the standard deviations of the  $\Phi$  and  $\Psi$  angles of the best matches in the TALOS-N database. For residue K72, there is no consensus among the database matches and, hence, no predicted torsion angles are shown. Secondary structure predictions are based on observed chemical shifts and classified as helix (H), strand (E), or loop (L). Predictions with a confidence > 0.75 are highly reliable. Based on sequence information, the secondary structure of I68 is classified as loop. Based on observed chemical shifts, the secondary structure of K72 is classified as loop. Torsion angles (black dots) of the AlphaFold-Multimer prediction (Figure 2.5c in the main manuscript) are also plotted for comparison. The horizontal solid and dashed black lines indicate the residues classified as  $\beta$ -strand by ChimeraX and VADAR, respectively.



### 3 Chapter 3 Co-sedimentation as the key to the investigation of wild-type FAT10 by MAS NMR spectroscopy

#### 3.1 Abstract

The ubiquitin-like modifier FAT10 targets substrate proteins for proteasomal degradation in an ubiquitin- and p97-independent manner. The protein consists of two ubiquitin-like domains (N- and C-domain), which are connected by a flexible linker. Its loosely folded and dynamic nature enables a fast and direct degradation that is even more pronounced in presence of the adapter protein NUB1L. Magic-angle spinning (MAS) NMR spectroscopy of isolated wild-type N-domain of human FAT10 results in spectra of poor quality as the protein aggregates. However, upon co-sedimentation with NUB1L, the spectra show excellent resolution allowing for the determination of the binding site within an otherwise disordered N-domain of FAT10. Thus, we propose sample preparation by co-sedimentation and the subsequent use of a MAS CryoProbe for enhanced sensitivity as a general strategy to study weakly folded proteins in small quantities by MAS NMR spectroscopy.

#### 3.2 Introduction

Proteins are the workhorses of cells.<sup>127</sup> They serve biological functions in virtually all cellular processes ranging from metabolic pathways, signal transduction to gene transcription.<sup>128</sup> In order to carry out their biological function, proteins were claimed to possess stably folded 3D structures under physiological conditions. However, with growing evidence for biologically active proteins that lack unique, well-defined equilibrium structures, the 'structure-function paradigm' had to be reassessed.<sup>129,130</sup> Rather, the native protein structure corresponds to a continuous spectrum of structural states from ordered to completely disordered and protein function can emerge from any state and structural transitions.<sup>131</sup> Similar to individual proteins, bound proteins can be described by a single 3D structure or show structural disorder or polymorphism in fuzzy complexes.<sup>114</sup> Considering structural studies of disordered proteins, the approach that must be taken is complicated, since the determination of a unique high-resolution structure is not attainable anymore.<sup>132</sup>

The ubiquitin-like modifier FAT10 (human leukocyte antigen-F adjacent transcript 10) functions as tag for proteasomal degradation in an ubiquitin-independent manner.<sup>69</sup> The protein's flexible, labile and dynamic nature is thought to enable a rapid and direct degradation.<sup>70,133</sup> FAT10, which consists of two ubiquitin-like domains (referred to as N- and C-domain) connected by a short linker,<sup>64</sup> is extremely short lived with an estimated cellular half-life of ~1 h and seems to be degraded along with its substrates.<sup>69</sup> Using differential scanning fluorimetry, the melting temperature of an N-terminally truncated variant of wild-type (WT) FAT10 was determined to 41 °C in contrast to 83 °C of ubiquitin.<sup>70</sup> MD simulations show that long-range salt bridges present in ubiquitin are absent rendering FAT10 flexible.<sup>133</sup> For reasons of

poor solubility and a tendency to aggregate, structural studies of FAT10 were complicated in the past.<sup>70,75</sup>

Upon deletion of seven N-terminal residues, Theng *et al.* (2014) could stabilize and determine the structure of the N-domain by solution-state NMR (PDB-ID 2MBE).<sup>75</sup> Heteronuclear single quantum coherence (HSQC) peaks of residues P59-T73 were not detected and the lack of slowly exchanging amide protons let them assume that the N-domain experiences global folding and unfolding.<sup>75</sup> By mutating the cysteine residues, Aichem *et al.* (2018) could express stabilized variants of full-length FAT10 and the individual domains.<sup>70</sup> In case of the cysteine-free (Cys-free) C-domain, structural characterization was possible by solution-state NMR (PDB-ID 6GF2). The structure of the Cys-free N-domain (N-FAT10-C0) was revealed by X-ray crystallography (PDB-ID 6GF1).<sup>70</sup> Although unaffected conjugation to substrate proteins (so-called FAT10ylation), Cys-free mutants of FAT10 and respective conjugates were degraded at slower rates by the 26S proteasome when compared to the WT protein and conjugates.<sup>70</sup> In presence of the adapter protein NUB1 (NEDD8 ultimate buster 1) and its longer isoform NUB1L, degradation of FAT10 and conjugates gets accelerated.<sup>69,72</sup> NUB1(L) comprises one N-terminal ubiquitin-like (UBL) domain and three C-terminal ubiquitin-associated (UBA) domains. The UBL domain is able to bind to the proteasomal ubiquitin receptors Rpn1 and Rpn10,<sup>73</sup> whereas the UBA domains interact with the N-domain of FAT10<sup>74</sup>.

Most recently, Arkinson *et al.* (2025) uncovered that the intrinsic instability and spontaneous unfolding of the N-domain of FAT10 are critical for NUB1 to bind and trap an unfolded state for engagement and rapid degradation by the 26S proteasome.<sup>71</sup> During the whole process of proteasomal turnover, the complex formation between NUB1 and FAT10 is slow and rate limiting and a free N-terminus of FAT10 is needed for degradation. NUB1 behaves like an ATP-independent chaperone for an unfolded N-domain of FAT10 that functions as degradation initiation region without the need for substrate unfolding by the segregase p97.<sup>71</sup> In our group, we used magic-angle spinning (MAS) NMR spectroscopy to study the complex of N-FAT10-C0 and NUB1L.<sup>134</sup> Based on carbon and nitrogen chemical shift assignments, TALOS-N<sup>100</sup> prediction of secondary structure and AlphaFold-Multimer<sup>93</sup> modelling, we concluded that NUB1L and partially unfolded N-FAT10-C0 interact *via* an intermolecular, anti-parallel  $\beta$ -sheet and *via* electrostatic and hydrophobic interactions forming a fuzzy complex.<sup>134</sup>

MAS NMR spectroscopy neither requires the system of interest to be soluble nor crystallizable<sup>43,135,136</sup> allowing for structural insights into otherwise inaccessible systems such as amyloid fibrils,<sup>137,138</sup> membrane proteins<sup>139,140</sup> or viral assemblies<sup>141,142</sup>. Moreover, MAS NMR spectroscopy is capable to explore dynamic behaviour at different time scales, chemical processes and reactivity as well as ligand binding and macromolecular interactions with atomic resolution.<sup>11</sup> A critical step for MAS NMR studies includes the preparation and thereby associated quality of the sample. Generally, micro- and nanocrystalline protein samples produced by batch crystallization or controlled precipitation show a superior chemical shift resolution to lyophilized powders.<sup>103</sup> In place of extended searches for suitable crystallization/precipitation conditions, proteins of sufficient molecular weight can be sedimented out of solution to obtain highly resolved MAS NMR spectra.<sup>143</sup> This was shown in the sedimented state, e.g. for the 480 kDa complex of 24 molecules of apoferritin<sup>143</sup> or dodecameric DnaB helicase (12x 59 kDa)<sup>144</sup>. In several other studies, co-sedimentation by ultracentrifugation was described as simple

method to prepare samples of protein-ligand or protein-protein complexes. In case of sedimented hepatitis B virus capsids containing different capsid assembly modulator drugs, the induced abnormal nucleocapsid assembly could be revealed on a molecular level.<sup>142</sup> To investigate the interactions between the DnaB helicase and the C-terminal domain of the corresponding DnaG primase of *Helicobacter pylori*, the complex was prepared by co-sedimentation of the two proteins directly into the MAS rotor.<sup>92</sup>

Here, we report the investigation of the WT N-domain of human FAT10 (N-FAT10-WT) by MAS NMR spectroscopy. Upon centrifugal concentration, isolated N-FAT10-WT starts to aggregate and corresponding spectra are of poor quality. In order to still get information on N-FAT10-WT, we take advantage of the known interaction with NUB1L. We succeed in the preparation of an N-FAT10-WT sample by co-sedimentation with NUB1L. Intriguingly, the resolution of the co-sedimented sample is excellent indicating that N-FAT10-WT has (in part) adopted a regular structure. As a strategy to deal with the insufficient signal-to-noise ratio (SNR) (the rotor is mostly filled with a 7-times bigger NUB1L protein compared to N-FAT10-WT), we explore the abilities of a Bruker MAS CryoProbe at 14.1 T and non-uniform sampling (NUS). A stretch of residues corresponding to I68-V81 within N-FAT10-WT is sequentially assigned. As for N-FAT10-C0 in complex with NUB1L, we observe again the N-terminus, even though the first residue is this time alanine and not glycine. This provides a proof that we see the N-terminus raising the question whether the N-terminus of N-FAT10 can bind to NUB1L regardless of the residue type. Apart from that, no further differences between the two samples can be identified: N-FAT10 interacts with NUB1L forming an intermolecular, anti-parallel  $\beta$ -sheet, while the remaining part of the N-domain is (statically or dynamically) disordered.

### 3.3 Materials and methods

Below we give short accounts of all procedures. Further specifics can be found in the Supporting Information.

**Cloning.** Molecular cloning was achieved by the classical restriction enzyme cloning approach using BsaI and XhoI restriction enzymes. WT N-domain of human FAT10 (amino acids 2-86) and an N-terminally truncated variant thereof (amino acids 5-86) were cloned into pSUMO vectors to express His<sub>6</sub>-SUMO-fusion proteins. Isolated candidate clones were verified by DNA sequencing.

**Expression and Purification of U-<sup>13</sup>C, <sup>15</sup>N-N-FAT10-WT.** WT N-domain of human FAT10 (comprising either amino acids 2-86 or 5-86) was expressed in *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells (Agilent Technologies). For uniform <sup>13</sup>C and <sup>15</sup>N labelling (U-<sup>13</sup>C, <sup>15</sup>N), bacteria cells were grown in M9 minimal medium supplemented with U-<sup>13</sup>C<sub>6</sub>-D-glucose and <sup>15</sup>N-ammonium chloride as exclusive sources of carbon and nitrogen. The purification (see Figure S3.1) was adapted from a previously published protocol.<sup>70</sup> Briefly, bacterial cells were lysed and the supernatant was applied to nickel affinity chromatography. After buffer exchange, Ulp1 protease was added for cleavage of the purification tag. The Ulp1 protease recognizes the 3D fold of the SUMO-tag preserving the sequence that starts in both cases with an alanine residue. The His-tagged protease and respective cleavage byproducts were

separated by a second nickel affinity chromatography. Further purification was achieved by size-exclusion chromatography yielding up to 3 mg per liter of M9 medium.

**Expression and Purification NUB1L.** Plasmid pSUMO coding for human NUB1L (amino acids 2-615) was kindly provided by the Institute of Cell Biology and Immunology Thurgau (Kreuzlingen, Switzerland) and used to express the His<sub>6</sub>-SUMO-fusion protein in *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells. To produce natural abundance (NA) NUB1L, bacteria cells were grown in LB medium. Purification required the same steps as for U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT. Yield was up to 15 mg per liter of LB medium.

**N-FAT10-WT.** During centrifugal concentration of U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT (amino acids 2-86), the protein started to aggregate. The white aggregate was packed into a MAS rotor by ultracentrifugation.

**Complex of N-FAT10-WT and NUB1L.** Co-sedimentation by ultracentrifugation, in a molar ratio of 1:1, was used to form the complex of NA NUB1L and U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT (amino acids 5-86) in the MAS rotor. In case of packing of the Bruker MAS CryoProbe rotor, the complex was co-sedimented into a polypropylene ultracentrifuge tube and transferred into the rotor by centrifugation using a swinging-bucket benchtop centrifuge. Teflon spacers were placed at the bottom and the top of the CryoProbe rotor.

**MAS NMR Spectroscopy.** MAS NMR measurements were performed at 18.8 T (<sup>1</sup>H Larmor frequency of 800.3 MHz) on a Bruker Avance NEO spectrometer equipped with a 3.2 mm E-free HCN Bruker MAS probe and a Bruker BCU II cooling unit. Additional measurements were conducted at 14.1 T (<sup>1</sup>H Larmor frequency of 600.3 MHz) using the Bruker MAS CryoProbe. The spinning frequency was 12.0 or 14.5 kHz. The sample temperature was 4 °C. Chemical shifts of <sup>13</sup>C are referenced to DSS in D<sub>2</sub>O (0.5 % by weight), chemical shifts of <sup>15</sup>N are referenced to liquid ammonia at 25 °C.

Sample integrity was verified by <sup>1</sup>H-<sup>13</sup>C and <sup>1</sup>H-<sup>15</sup>N cross-polarization (CP) experiments as well as <sup>13</sup>C-<sup>13</sup>C correlation spectra recorded with the dipolar assisted rotational resonance (DARR)<sup>34</sup> pulse sequence. A mixing time of 10.0 ms was used in case of N-FAT10-WT (amino acids 5-86) in complex with NUB1L for the observation of mainly one- and two-bond correlations. From <sup>13</sup>C-<sup>13</sup>C correlation patterns, amino acid residue types can be identified, but sequential assignment is not possible. For this purpose, <sup>15</sup>N-<sup>13</sup>C spectra were recorded for N-FAT10-WT (amino acids 5-86) in complex with NUB1L with the z-filtered transferred echo double-resonance (ZF TEDOR)<sup>32</sup> pulse sequence or with SPECIFIC-CP<sup>44</sup> from backbone nitrogen atoms either to backbone C' or C<sub>α</sub> carbons. The TEDOR mixing time was set to ~1 ms to emphasize N-C<sub>α</sub> and N-C' magnetization transfers along the protein backbone and, in the side chains, to carbon nuclei directly bonded to nitrogen nuclei.

<sup>15</sup>N-<sup>13</sup>C-<sup>13</sup>C NCOX and NCACX spectra of N-FAT10-WT (amino acids 5-86) in complex with NUB1L were acquired with SPECIFIC-CP and COmbined R<sub>2</sub><sub>n</sub><sup>v</sup>-Driven (CORD) homonuclear carbon mixing<sup>95</sup> of 50 ms using NUS<sup>145,146</sup> of 25 %. For the reliable reconstruction of NUS data, a high SNR is needed. To improve the SNR, a second set of <sup>15</sup>N-<sup>13</sup>C-<sup>13</sup>C NCOX and NCACX spectra was recorded and added to the first one resulting in NUS of 50 %.

An overview of all MAS NMR experiments and acquisition parameters is given in Tables S3.1 and S3.2. Spectra were processed in NMRPipe<sup>96</sup>. 3D NUS data were reconstructed with the Iterative Soft

Thresholding (IST) method<sup>147</sup> as implemented in NMRPipe. The CcpNmr Analysis V3 software was used for the assignment of protein resonances.<sup>97</sup>

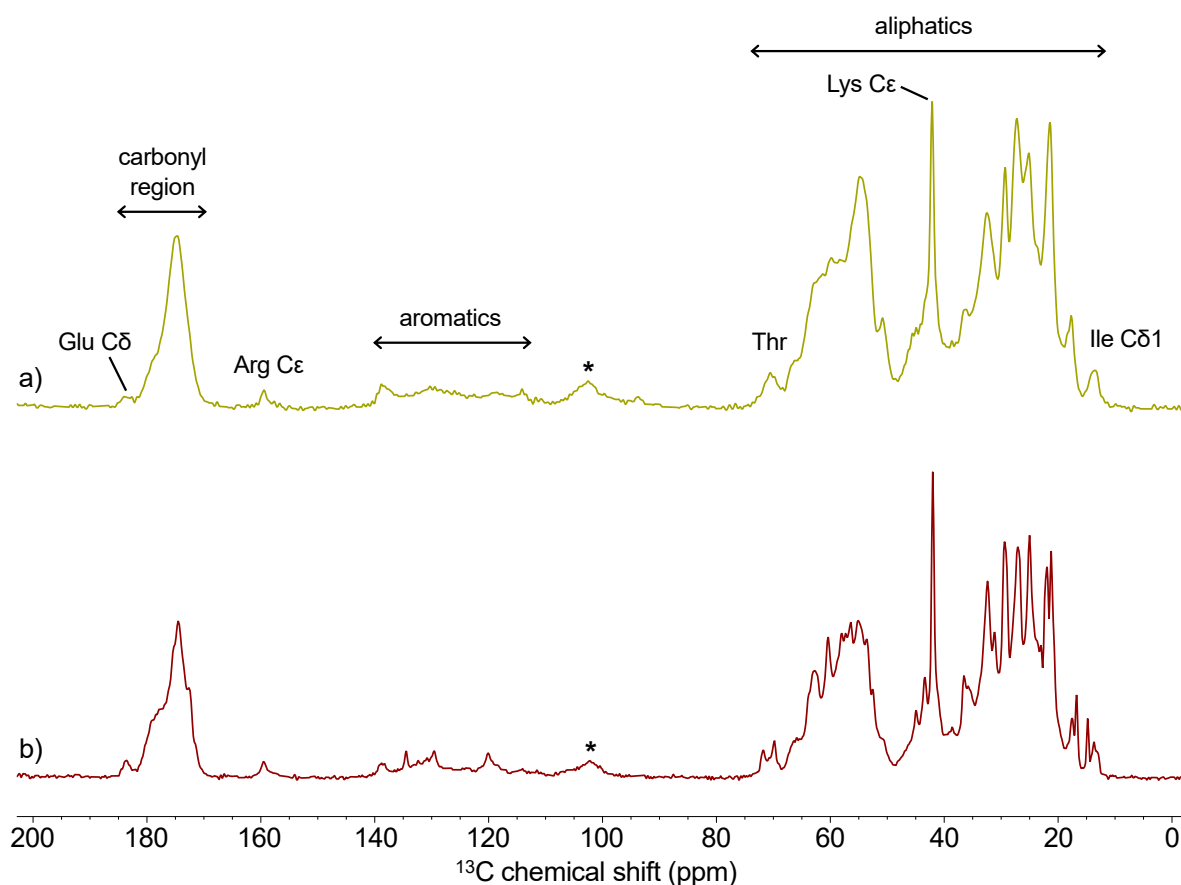
**Torsion Angle Analysis.** Backbone torsion angles  $\Phi$  and  $\Psi$  and secondary structure were predicted empirically based on chemical shifts of assigned N, C', C $\alpha$  and C $\beta$  nuclei.<sup>98,99</sup> For this purpose, we used the program TALOS-N,<sup>100</sup> which combines a set of trained neural networks with efficient mining of a database of proteins.

**AlphaFold-Multimer Modelling.** AlphaFold<sup>148</sup> v2.3.2 was downloaded from Github (<https://github.com/google-deepmind/alphafold/tree/v2.3.2>) and installed on the local Scientific Compute Cluster of the University of Konstanz (SCCKN). To run AlphaFold in the multimer mode,<sup>93</sup> a multi-sequence FASTA file including the WT N-domain of FAT10 (amino acids 5-86) and NUB1L (amino acids 1-615) was provided. Database presets were set to reduced and a maximum template release date of 2004-11-28 was used to exclude structural templates of FAT10 and NUB1L deposited in the PDB. By default, the multimer system ran 5 seeds per model resulting in 25 total predictions. The 25 predictions were ranked by model confidence using a weighted combination of predicted TM (pTM) and Interface pTM (ipTM) scores (model confidence = 0.8·ipTM + 0.2·pTM).<sup>93</sup> Only the best ranked model was relaxed. Structure visualization was performed with UCSF ChimeraX.<sup>101</sup>

### 3.4 Results

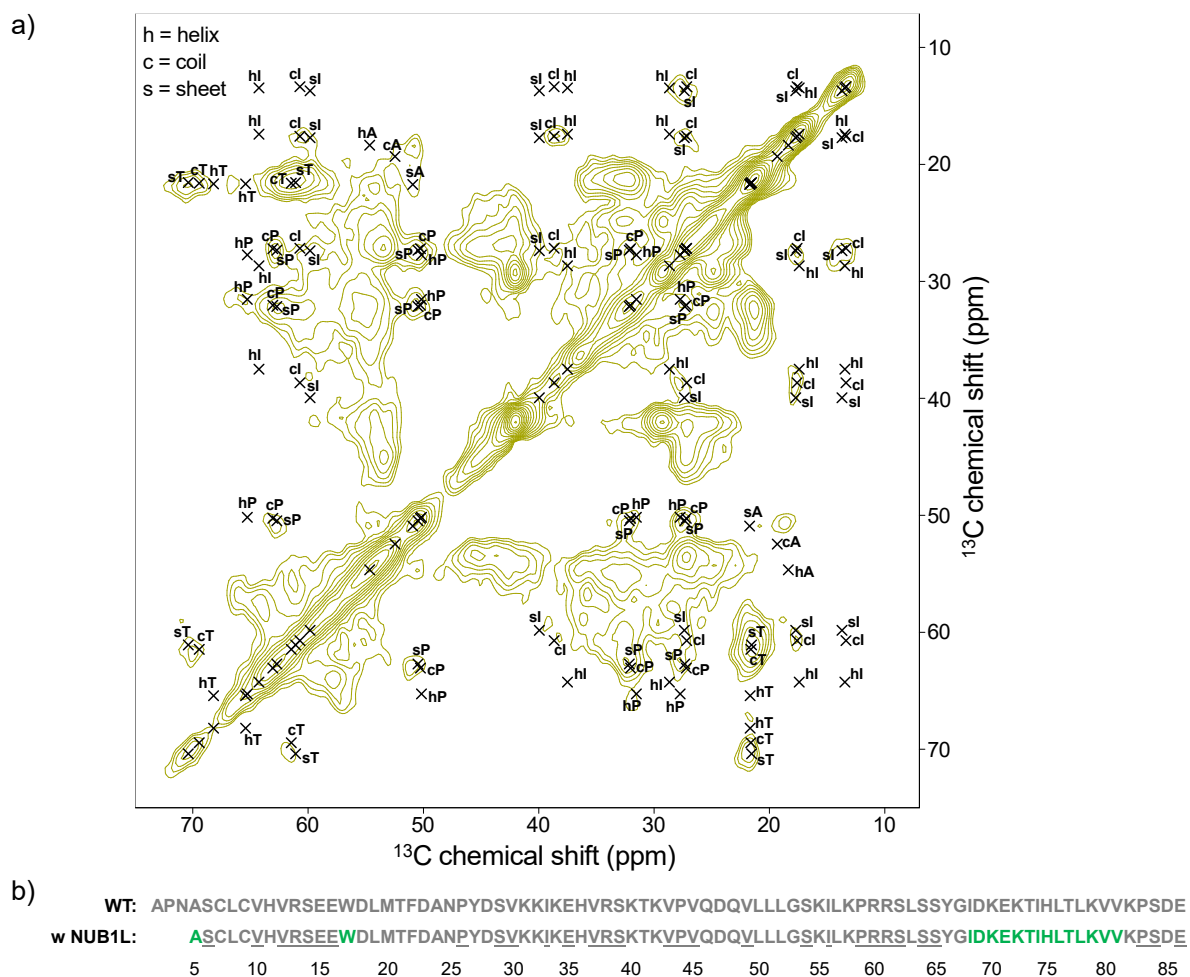
In contrast to the Cys-free mutant of the N-domain of FAT10, for which concentrations up to 25 mg/mL could be reached and microcrystals could be grown, the WT N-domain of FAT10 (amino acids 2-86) was unstable in solution and started to aggregate upon centrifugal concentration. To check for the sample condition, we sedimented the aggregated N-FAT10-WT (amino acids 2-86) into a MAS rotor by ultracentrifugation. <sup>1</sup>H-<sup>13</sup>C and <sup>1</sup>H-<sup>15</sup>N CP as well as <sup>13</sup>C-<sup>13</sup>C correlation spectra measured at 18.8 T are shown in Figures 3.1a, 3.2a and S3.2a.

The <sup>1</sup>H-<sup>13</sup>C CP spectrum of N-FAT10-WT (amino acids 2-86) in Figure 3.1a features low signal dispersion and broad lines with linewidths  $\geq 1$  ppm. The poor spectral resolution is most notably in the carbonyl and C $\alpha$  carbon regions (170-185 ppm and 50-68 ppm, respectively). Isolated narrow signals at the edges of the aliphatic region, typical for Thr C $\alpha$ /C $\beta$  and Ile C $\delta 1$  atoms in confined positions, are absent indicative for a loss in structural organization. As reported for ubiquitin in lyophilized and microcrystalline states, nitrogen atoms are even more sensitive to the sample condition compared to carbon atoms.<sup>149</sup> The respective <sup>1</sup>H-<sup>15</sup>N CP spectrum of N-FAT10-WT (amino acids 2-86) in Figure S3.2a shows severe line broadening with linewidths  $> 1$  ppm and is characterized by widespread peaks of backbone nitrogen atoms around 120 ppm as well as of side chain nitrogens that can be assigned to arginine (68-87 ppm) and lysine (33 ppm) residues.



**Figure 3.1:**  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization spectra of (a) N-FAT10-WT (amino acids 2-86) and (b) N-FAT10-WT (amino acids 5-86) in complex with NUB1L. Spectra were recorded at 18.8 T with a spinning frequency of 14.5 kHz and a temperature of approximately 4 °C. The stars indicate spinning sidebands.

The measurement of a 2D spectrum in Figure 3.2a confirms the poor spectral resolution as observed in the 1D spectra. Based on chemical shift statistics from deposited solid-state NMR data in the Biological Magnetic Resonance Bank (BMRB),<sup>150</sup> we plotted the average carbon chemical shifts typical for the helical, coil or sheet secondary structure of different amino acids (alanine, isoleucine, proline and threonine) as black crosses together with the contours of the 2D spectrum. Regardless of amino acid type, the helical conformation stands out, whereas coil and sheet conformations show similar average values. Moreover, as part of the backbone, the  $\text{C}_\alpha$  atoms experience the largest perturbations. With respect to chemical shifts of N-FAT10-WT (amino acids 2-86), alanine, isoleucine, proline and threonine residues rather tend to be in coil or sheet than in helical conformation. The broad signals are a sum of different conformations and environments allowing for the identification of amino acid types, but making a site-specific assignment of resonances impossible. Taken together, these observations let assume that N-FAT10-WT (amino acids 2-86) has lost part or most of its 3D organization in the aggregated form, with some elements retaining secondary structure.



**Figure 3.2:** (a)  $^{13}\text{C}$ - $^{13}\text{C}$  DARR spectrum with a mixing time of 50 ms of N-FAT10-WT (amino acids 2-86) measured at 18.8 T, a spinning frequency of 14.5 kHz and a temperature of approximately 4 °C. Average chemical shifts of alanine, isoleucine, proline and threonine in helix (h), coil (c) or sheet (s) conformation<sup>150</sup> are included and labelled accordingly (shown as black crosses). (b) Overview of sequential assignments of N-FAT10-WT (amino acids 2-86) and N-FAT10-WT (amino acids 5-86) in complex with NUB1L. Green = unambiguous and grey = unobserved. Of the grey, underlined residues of N-FAT10-WT (amino acids 5-86) in complex with NUB1L, we observe 1xE, 1xI, 1xP, 1xR, 2xS and 1xV, but sequential assignment was not possible.

With a molecular weight smaller than 10 kDa, sedimentation by ultracentrifugation out of a diluted solution is not feasible for N-FAT10-WT. In case of ubiquitin, sedimentation was just possible as the protein can be obtained in high concentrations and forms non-covalent dimers.<sup>151,152</sup> To overcome this problem and gather information on the WT N-domain of FAT10, we exploited the alternative route of co-sedimentation with the known interaction partner NUB1L.

Upon addition of the 7-times bigger NUB1L in a molar ratio of 1:1, an N-FAT10-WT sample could be prepared by co-sedimentation without the needs for high concentrations avoiding aggregation. As we could establish complex formation for the stabilized, Cys-free N-domain of FAT10 comprising amino acids 5-86 and NUB1L in our previous work,<sup>134</sup> we used the N-terminally shortened variant of WT N-FAT10 for co-sedimentation. In this way, the identification of potential differences in binding between the two complexes is not complicated by differences in sequence length.

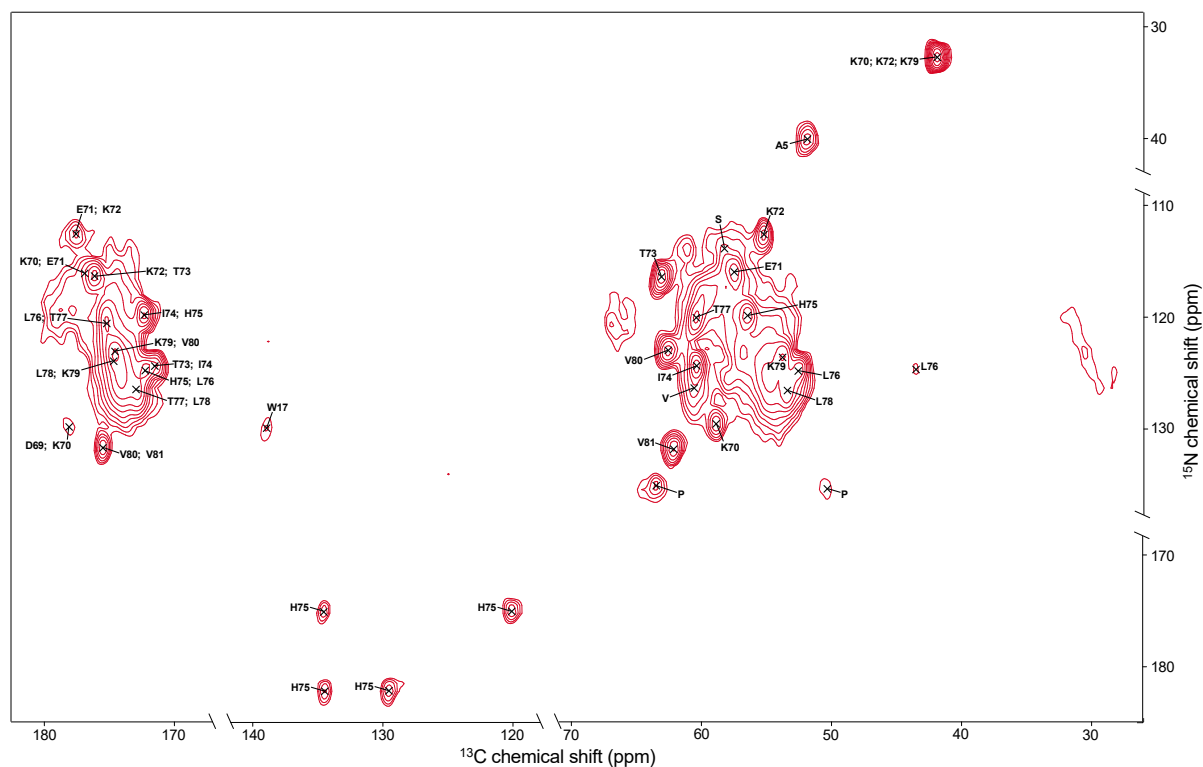
Once sedimented into the MAS rotor, the complex is stable over the course of months, when stored at 4 °C in the fridge, having a yellowish, semi-transparent appearance with a sticky texture. In a first



The  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectrum in Figure 3.3 contains one-, two- and some three-bond correlations. In the  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectrum in Figure 3.4, mostly one-bond  $\text{N}-\text{C}_\alpha$  and  $\text{N}-\text{C}'$  as well as correlations between side chain nitrogen and carbon atoms of lysine, arginine, tryptophane and histidine residues can be observed. Corresponding  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectra recorded at a lower magnetic field of 14.1 T using the MAS CryoProbe show the same spectral features. Differences in the  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectra, including the observation of correlations that belong to the N-terminus, a proline and the side chains of lysine, arginine, tryptophane and histidine residues (compare Figures 3.4 and S3.4), can be explained by the use of alternative pulse sequences and different settings of spectral widths in the nitrogen dimension.

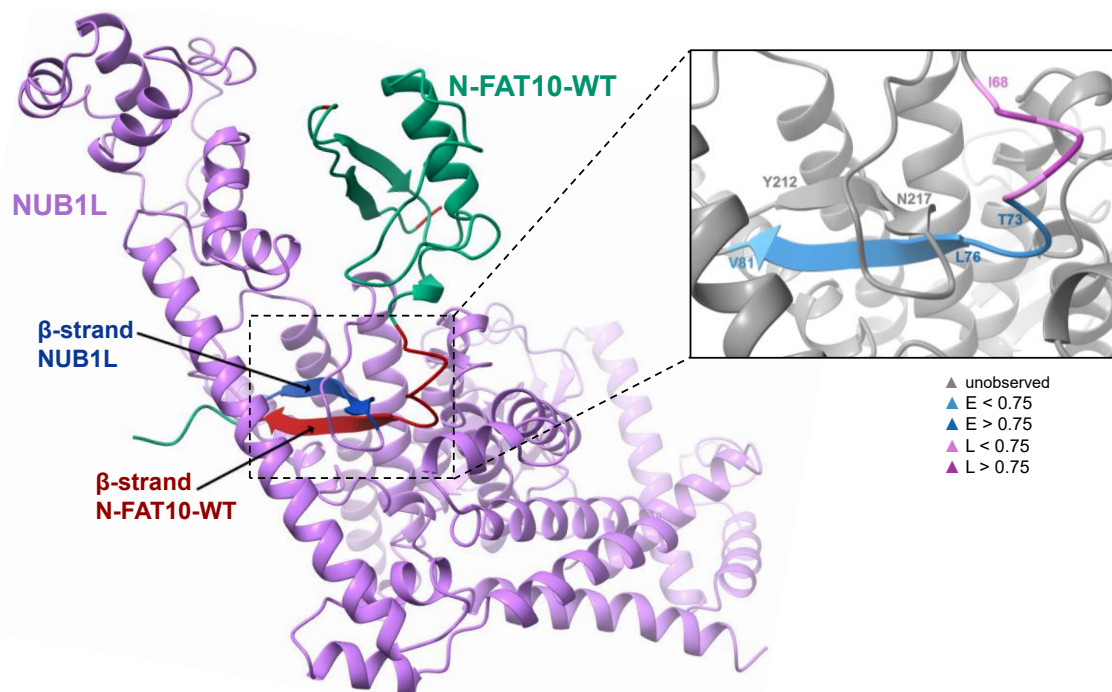
In comparison to the  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectrum of N-FAT10-WT (amino acids 2-86) with throughout broad peaks (Figure 3.2a), well-resolved peaks can be observed for the co-sedimented sample (see Figures 3.3 and 3.4) that can be assigned to residues A5, W17, I68-V81 and seven residues with missing backbone connectivities (1xE, 1xI, 1xP, 1xR, 2xS and 1xV) suggesting that the remaining part of N-FAT10-WT (amino acids 5-86) is disordered. Interestingly, spectra of the two complexes containing either Cys-free N-FAT10 or WT N-FAT10 and NUB1L almost perfectly overlay with the exception of resonances that can be assigned to N-terminal residues. The Cys-free N-domain of FAT10 bears an additional glycine residue at the N-terminus, whereas the WT N-domain has an N-terminal alanine residue.

Correlations within the  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra were mostly limited to backbone and  $\text{C}_\beta$  carbons emphasizing the importance of the  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{13}\text{C}$ - $^{15}\text{N}$  correlation spectra confirming the site-specific assignment of resonances.



**Figure 3.4:**  $^{15}\text{N}$ - $^{13}\text{C}$  ZF TEDOR spectrum of N-FAT10-WT (amino acids 5-86) in complex with NUB1L with a mixing time of 1.1 ms acquired at 18.8 T with a spinning frequency of 14.5 kHz and a temperature of approximately 4 °C. Peak symbols of site-specifically or residue type assigned resonances are included and labelled accordingly. Cross peaks from side chains of Arg ( $^{15}\text{N}$  chemical shift 67-90 ppm) are also observed but not plotted here.

TALOS-N secondary structure analysis based on backbone N, C' and C $\alpha$  as well as C $\beta$  chemical shifts and sequence information predicts residues I68-K72 to be in coil conformation and residues T73-V81 to form a  $\beta$ -strand (Figures 3.5 inset and S3.5). As site-specific assignments were only possible for I68-V81 (except for A5 and W17), it seems most likely that these residues are involved in binding to NUB1L. Consulting AlphaFold-Multimer for the prediction of the interaction of N-FAT10-WT (amino acids 5-86) and NUB1L, the best ranked structure (model confidence of 0.77) shows an opened FAT10 N-domain that forms an intermolecular, anti-parallel  $\beta$ -sheet with NUB1L (Figure 3.5).



**Figure 3.5:** AlphaFold-Multimer prediction for the interaction of N-FAT10-WT (amino acids 5-86) and NUB1L. The residues of N-FAT10-WT that are observed by MAS NMR are highlighted in red. The inset shows the secondary structure prediction for residues I68-V81 projected on the AlphaFold-Multimer structure (*E* stands for strand and *L* for loop). See Figure S3.5 for plots of the backbone torsion angles and further specifics.

### 3.5 Discussion

The sedimented sample of N-FAT10-WT (amino acids 2-86) was of poor quality and corresponding MAS NMR spectra showed low resolution and no good signal dispersion (Figures 3.1a, 3.2a and S3.2a), typical for a protein of structural heterogeneity. Molecular dynamics (MD) simulations pinpoint high flexibility in stretches of residues 40-50 and 70-80 for the WT N-domain of FAT10.<sup>70</sup> Mutation of the cysteine residues stabilizes the N-domain of FAT10.<sup>70</sup> For the highly dynamic and flexible, less stable N-FAT10-WT, the free energy barrier for (partial) unfolding is lowered<sup>133</sup> compared to the stabilized, Cys-free mutant. Particularly the (partially) unfolded states are prone to aggregation.<sup>153</sup> Upon concentration, the probability of (partially) unfolded molecules to collide with each other is increased leading to aggregation.

In a dilute solution, the probability of two (partially) unfolded N-FAT10-WT molecules to collide is reduced and aggregation is circumvented. However, sedimentation of proteins with a molecular weight smaller than 10 kDa requires high concentrations.<sup>18</sup> To overcome this limitation, we added NUB1L in an equimolar ratio to a dilute solution of N-FAT10-WT (amino acids 5-86). NUB1L is a known interaction

partner<sup>72,74</sup> and has the basic holdase function of a chaperone,<sup>153</sup> i.e. it is able to bind the N-domain of FAT10 in a partially unfolded state. In this way, we could prepare a MAS NMR sample simply co-sedimenting N-FAT10-WT (amino acids 5-86) and NUB1L by ultracentrifugation.

Just recently, we could determine by MAS NMR spectroscopy residues T73-V81 within N-FAT10-C0 essential for binding to NUB1L forming an intermolecular  $\beta$ -sheet. Much to our surprise, we found that MAS NMR spectra of both complex preparations can be overlaid almost perfectly with the exception of resonances that can be assigned to the very N-terminal residues. Lysines comprise a free amine group in their side chains with chemical shifts of  $\sim 30$  ppm for the nitrogen and  $\sim 40$  ppm for the neighboring  $C_\epsilon$  atom. In the  $^{15}\text{N}$ - $^{13}\text{C}$  ZF TEDOR spectra, we could observe upfield-shifted peaks in terms of the nitrogen dimension in a region typical for lysines or the free amine of the N-terminus.<sup>106</sup> With carbon chemical shifts that match the  $C_\alpha$  atoms of glycine or alanine rather than the  $C_\epsilon$  atom of a lysine, we assigned the resonances to the N-terminal residues. Corresponding spin system patterns of a glycine or an alanine could be detected in the  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra.

The number of resonances, that could be assigned site-specifically, was limited and corresponds alongside the N-terminal alanine and W17 (there is only one tryptophane present in the sequence of the N-domain) to a stretch of residues I68-V81 within N-FAT10-WT (amino acids 5-86) (see Figures 3.3, 3.4, S3.3 and S3.4). For further seven residues (1xE, 1xI, 1xP, 1xR, 2xS and 1xV) side chain patterns are evident, but their backbone connectivities could not be observed. The generally reduced number of signals with respect to the length of the sequence let assume that the remaining part of the protein is disordered. TALOS-N analysis of assigned chemical shifts predicts residues T73-V81 in  $\beta$ -strand conformation. Together with AlphaFold-Multimer modelling (Figure 3.5), results can be interpreted such that partially unfolded N-FAT10-WT (amino acids 5-86) interacts with NUB1L *via* an intermolecular  $\beta$ -sheet and *via* electrostatic and hydrophobic interactions at the scattered, non-sequentially assigned residues forming a fuzzy complex<sup>114</sup>. Hence, the WT N-domain behaves just like the Cys-free N-domain of FAT10 when bound to NUB1L.

At the site of the 26S proteasome, FAT10 and its conjugates are degraded in a p97-independent manner.<sup>70</sup> For ubiquitin-mediated degradation, tightly folded substrates have to be unfolded by the segregase p97 (also known as VCP) as an unstructured initiation region is needed for engagement and activation of the proteasome.<sup>154-156</sup> In FAT10-mediated degradation, the unstructured initiation region is already present in form of the partially unfolded N-domain. N- and C-terminal residues are usually flexible and thus undetected by MAS NMR spectroscopy. As the N-terminal alanine of N-FAT10-WT (amino acids 5-86) can be observed and assigned specifically, it seems likely that NUB1L interacts with and pre-positions the N-terminus for a fast engagement and subsequent proteasomal degradation of FAT10 and conjugated substrate proteins. According to our data, Cys-free N-FAT10 and N-FAT10-WT interact in the same way with NUB1L. The difference in proteasomal degradation speeds can be possibly explained by different complex formation rates: The highly dynamic and flexible WT protein unfolds more easily and consequently, complex formation and thereby associated proteasomal degradation are accelerated compared to the stabilized, Cys-free mutant.

We think that the demonstrated co-sedimentation approach with a known interaction partner might be generally applicable to unstable, loosely folded proteins. Upon simple addition of the known interaction

partner to a diluted solution of the “hard-to-handle” protein, a co-sedimented sample for MAS NMR spectroscopy can be obtained by ultracentrifugation to guarantee for sufficient hydration at the same time. From our experiences, it is worth to spend time on the design of the rotor packing tools to avoid loss of material and repeated sample preparations that can be time-consuming and expensive. A collection of different designs for 3D-printable centrifugal devices is reported by Osborn Popp *et al.* (2023).<sup>157</sup> One possible approach for the combination of two samples to enhance sensitivity includes the following steps: The first preparation can be directly packed into a large volume MAS rotor by ultracentrifugation. The second preparation is pelleted into an ultracentrifugal polypropylene tube. Subsequently, the pelleted sample can be transferred into the MAS rotor containing the first preparation under low RCF using a packing tool fitting the rotor on the one side and the polypropylene tube on the other side.

In this work, we show how sample preparation involving co-sedimentation with NUB1L can increase the resolution of MAS spectra of an intrinsically unstable and loosely folded WT N-domain of FAT10. With the enhanced sensitivity of the Bruker MAS CryoProbe and a NUS acquisition scheme it was possible to measure <sup>15</sup>N-<sup>13</sup>C-<sup>13</sup>C correlation spectra. Site-specific assignments correspond to the N-terminal alanine, W17 and residues I68-V81 within N-FAT10-WT (amino acids 5-86). NUB1L and N-FAT10-WT (amino acids 5-86) form a fuzzy complex and interact *via* an intermolecular  $\beta$ -sheet. The N-terminal residue is very likely bound to NUB1L, so that the otherwise disordered N-domain of FAT10 can be presented to the 26S proteasome for a rapid and direct degradation.

### 3.6 Conclusion

Isolated N-FAT10-WT started to aggregate even at low concentrations and corresponding spectra were characterized by broad lines typical for unfolded proteins displaying structural heterogeneity. Upon co-sedimentation with the adapter protein NUB1L, highly resolved spectra could be acquired for N-FAT10-WT. The easy implementation of co-sedimentation with a known binding partner is suggested to be a general strategy to access weakly folded proteins by MAS NMR spectroscopy. The complexes of NUB1L with either N-FAT10-C0 or N-FAT10-WT have the same or a very similar structure. They form an intermolecular  $\beta$ -sheet and the N-terminus of the N-domain of FAT10 is bound to NUB1L regardless of the residue type.

## 3.7 Supporting Information

### 3.7.1 Protein expression and purification

#### **U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT**

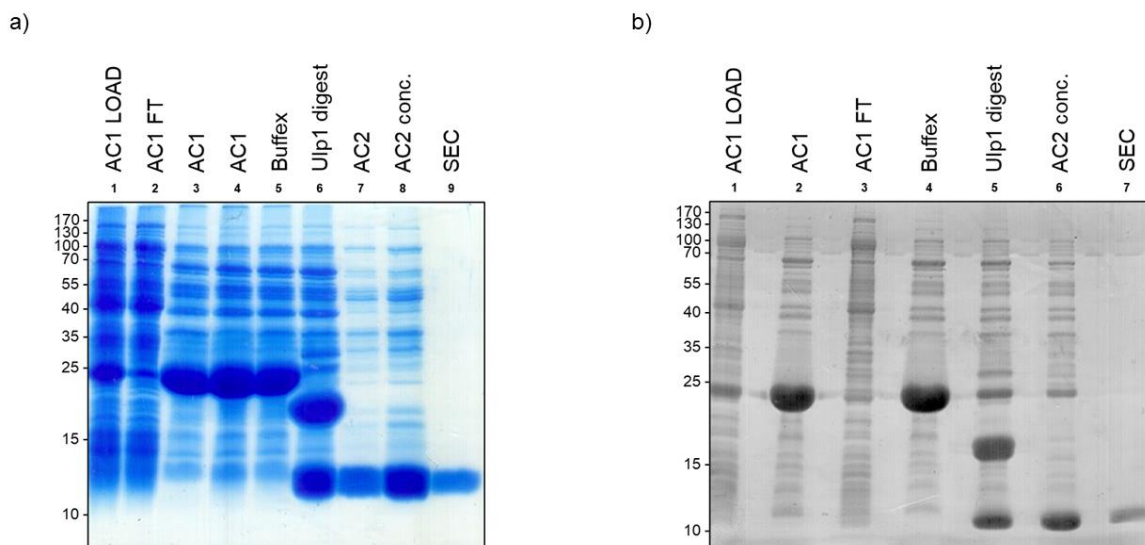
Wild-type N-domain of human FAT10 (amino acids 2-86) and an N-terminally truncated variant thereof (amino acids 5-86) were expressed as His<sub>6</sub>-SUMO-fusion proteins in *E. coli* BL21-CodonPlus(DE3)-RIPL competent cells (Agilent Technologies). For uniform <sup>13</sup>C and <sup>15</sup>N labelling, 3.6 g of U-<sup>13</sup>C<sub>6</sub>-D-glucose and 0.5 g of <sup>15</sup>N-ammonium chloride were added per liter of M9 minimal medium. Bacteria cells were grown at 37 °C to an OD<sub>600</sub> of 0.5-0.6, induced with 0.4 mM IPTG at 21 °C overnight and harvested by centrifugation (RCF 4000 g, 10 min, 8 °C). Harvested cells were lysed in lysis buffer (20 mM TRIS-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole, 1 mM TCEP, 10 % v/v glycerol, 0.1 % v/v Triton X-100, 100 µg/mL lysozyme, 1 mM PMSF and EDTA-free protease inhibitor). After sonication, cell debris was removed by centrifugation (RCF 47000 g, 30 min, 8 °C). The supernatant was filtered and applied to Ni-NTA Agarose beads (Macherey-Nagel) for affinity chromatography. The protein was eluted with elution buffer (20 mM TRIS-HCl (pH 8.0), 300 mM NaCl, 500 mM imidazole and 1 mM TCEP) and buffer exchanged to binding buffer (20 mM TRIS-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole and 1 mM TCEP). His-tagged Ulp1 protease was allowed to act overnight (4 µg of protease per 1 mg of recombinant protein). The Ulp1 protease recognizes the 3-dimensional fold of the SUMO-tag preserving the sequence that starts in both cases with an alanine residue. After cleavage, His-tagged Ulp1 protease and cleavage byproducts were separated by a second Ni-affinity chromatography using binding buffer. The volume of the flow-through containing U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT (amino acids 2-86 or 5-86) was reduced to 5 mL and filtered. Further purification was achieved by size-exclusion chromatography (Cytiva, formerly Amersham Biosciences, ÄKTA pure chromatography system equipped with a HiLoad 16/600 Superdex 75 pg column) yielding up to 3 mg per liter of M9 medium. Before preparation of the samples for MAS NMR, U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT (amino acids 2-86 or 5-86) was in 20 mM HEPES (pH 7.5), 150 mM NaCl and 1 mM TCEP and concentrated to specified values (Amicon Ultra-15 centrifugal filter unit, MWCO 3 kDa).

#### **NUB1L**

Human NUB1L (amino acids 2-615) was expressed as a His<sub>6</sub>-SUMO-fusion protein also in *E. coli* BL21CodonPlus(DE3)-RIPL competent cells. The pSUMO plasmid was kindly provided by the Institute of Cell Biology and Immunology Thurgau (Kreuzlingen, Switzerland). Bacteria cells were grown in LB medium at 37 °C to an OD<sub>600</sub> of 0.7, induced with 0.4 mM IPTG at 21 °C overnight and harvested by centrifugation (RCF 4000 g, 10 min, 8 °C). The harvested cells were lysed in lysis buffer (20 mM TRIS-HCl (pH 7.5), 150 mM NaCl, 20 mM imidazole, 1 mM TCEP, 10 % v/v glycerol, 0.1 % v/v Triton X-100, 100 µg/mL lysozyme, 1 mM PMSF and EDTA-free protease inhibitor). After sonication, cell debris was removed by centrifugation (RCF 27000 g, 30 min, 8 °C). The supernatant was filtered and applied to Ni-NTA Agarose beads for affinity chromatography. The protein was eluted with elution buffer (20 mM TRIS-HCl (pH 7.5), 150 mM NaCl, 500 mM imidazole and 1 mM TCEP) and buffer exchanged to binding buffer (20 mM TRIS-HCl (pH 7.5), 150 mM NaCl, 20 mM imidazole and 1 mM TCEP). His-

## Chapter 3 Co-sedimentation as the key to the investigation of wild-type FAT10 by MAS NMR spectroscopy

tagged Ulp1 protease was added to work overnight (4  $\mu\text{g}$  of protease per 1 mg of recombinant protein). After cleavage, His-tagged Ulp1 protease and cleavage byproducts were separated by a second Ni-affinity chromatography using binding buffer. The volume of the flow-through containing natural abundance NUB1L was reduced to 5 mL and filtered. Further purification was achieved by size-exclusion chromatography yielding up to 15 mg of pure NUB1L per liter of LB medium. To form the complex with N-FAT10-WT (amino acids 5-86), the protein was in 20 mM HEPES (pH 7.5), 150 mM NaCl and 1 mM TCEP, concentrated to approximately 20 mg/mL (Amicon Ultra-15 centrifugal filter unit, MWCO 30 kDa).



**Figure S3.1:** SDS-PAGE analysis for the purifications of (a) N-FAT10-WT (amino acids 2-86) and (b) N-FAT10-WT (amino acids 5-86). 5  $\mu\text{L}$  of gel sample buffer with SDS (reducing conditions) were added to 20  $\mu\text{L}$  of each sample. After boiling for 5 min at 95  $^{\circ}\text{C}$ , samples were loaded on a 15 % polyacrylamide gel and separated applying for 18-20 min a constant voltage of 120 V and for 55-65 min a constant voltage of 160 V followed by colloidal Coomassie staining. For molecular weight estimation, PageRuler Prestained Protein Ladder size standards (Thermo Fisher) were used. AC = Affinity Chromatography, FT = Flow-Through, Buffex = Buffer exchange, SEC = Size-Exclusion Chromatography.

### 3.7.2 Preparation of samples for MAS NMR

#### **N-FAT10-WT**

During centrifugal concentration of U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT (amino acids 2-86), the protein started to aggregate. Ultracentrifugation was used to assure dense packing of the white aggregate in a 3.2 mm MAS rotor. A home-made packing tool manufactured at the workshop of the University of Konstanz following Mandal *et al.* (2017)<sup>94</sup> containing the protein suspension was centrifuged for 4 h at RCF 210000 g and 4 °C. After the removal of supernatant, the rotor containing sedimented protein was closed with a Vespel cap.

#### **N-FAT10-WT in complex with NUB1L**

To prepare the complex for MAS NMR spectroscopy at 18.8 T, co-sedimentation by ultracentrifugation into the rotor was used. 3 mL of U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT (amino acids 5-86) at a concentration of 1 mg/mL was mixed with 1 mL of natural abundance NUB1L with a concentration of 20 mg/mL, leading to a molar ratio of 1:1. The packing tool containing this solution was centrifuged at RCF 210000 g and 4 °C. As the holding capacity of the packing tool is only 1 mL, the run was interrupted to remove supernatant and add more of the original solution over the course of several days. Packing efficiency was assessed by UV-Vis absorbance measurements of the supernatant at 280 nm. The 3.2 mm rotor was closed with a Vespel cap. For MAS spectroscopy at 14.1 T using the Bruker MAS CryoProbe, a polypropylene tube suitable for ultracentrifugation containing the solution of U-<sup>13</sup>C,<sup>15</sup>N-N-FAT10-WT (amino acids 5-86) and natural abundance NUB1L as described above was centrifuged for several days at RCF 210000 g and 4 °C. The supernatant was removed and the co-sedimented sample was transferred into the 3.2 mm rotor *via* centrifugation of the packing tool in a swinging-bucket benchtop centrifuge for 2 h at RCF 3100 g and 4 °C.

### 3.7.3 MAS NMR spectroscopy

MAS NMR experiments were performed at 18.8 T (<sup>1</sup>H Larmor frequency of 800.3 MHz) in the NMR Core Facility at the University of Konstanz and at 14.1 T (<sup>1</sup>H Larmor frequency of 600.3 MHz) at Bruker Biospin Fällanden. The console was Bruker Avance NEO. A Bruker 3.2 mm E-free HCN MAS probe or the MAS CryoProbe were used, together with regular-wall 3.2 mm ZrO rotors (maximum sample volume 32.1 µL) or CryoProbe thin-wall 3.2 mm silicon nitride rotors (maximum sample volume 90 µL). During experiments, the sample was kept at a temperature of approximately 4 °C. Calibration of the Bruker Cooling Unit (BCU) II was performed by monitoring the chemical shift of <sup>79</sup>Br in KBr powder.<sup>121</sup> The carbon dimension was referenced indirectly to DSS in D<sub>2</sub>O (0.5 % by weight),<sup>122</sup> i.e., the <sup>13</sup>C adamantane methylene peak is observed at 40.49 ppm. The nitrogen dimension was referenced to liquid ammonia at 25 °C.<sup>123</sup> Table S3.1 provides an overview of the experiments performed per sample.

**Table S3.1:** Summary of MAS NMR experiments.

Experiment	Number of scans	Non-uniform sampling	$^{15}\text{N}$ - $^{13}\text{C}$ / $^{13}\text{C}$ - $^{13}\text{C}$ mixing (ms)	Temperature ( $^{\circ}\text{C}$ )	MAS frequency (kHz)	Field (T)	Measurement time
<b><math>\text{U-}^{13}\text{C}, ^{15}\text{N}</math>-FAT10-WT (amino acids 2-86)</b>							
hC	512			4	14.5	18.8	
hN	256			4	14.5	18.8	
DARR	64		50	4	14.5	18.8	15.5 h
<b><math>\text{U-}^{13}\text{C}, ^{15}\text{N}</math>-FAT10-WT (amino acids 5-86) in complex with natural abundance NUB1L</b>							
hC	256			4	14.5	18.8	
hN	512			4	14.5	18.8	
DARR	128		10	4	14.5	18.8	3 d 0.5 h
ZF TEDOR	512		1.1	4	14.5	18.8	3 d 8 h
NCO	128 (2x)			4	12.0	14.1	3.5 h (2x)
NCA	128 (2x)			4	12.0	14.1	3.5 h (2x)
NCOCX	256 (2x)	25 %	50 (CORD)	4	12.0	14.1	3 d 4 h (2x)
NCACX	384/376	25 %	50 (CORD)	4	12.0	14.1	4 d 18.5 h/ 4 d 16 h

Typical  $\pi/2$ -pulse lengths were 2.8-3.0  $\mu\text{s}$  for  $^1\text{H}$ , 4.1-5.0  $\mu\text{s}$  for  $^{13}\text{C}$  and 6.0-9.0  $\mu\text{s}$  for  $^{15}\text{N}$ . The  $^1\text{H}$ - $^{13}\text{C}$ / $^1\text{H}$ - $^{15}\text{N}$  cross-polarization (CP)<sup>124</sup> step was realized using a linearly ramped radio frequency (RF) field on the  $^1\text{H}$  channel from 90 to 100 % amplitude; the centre of the ramp was set to match the  $n = +1$  Hartmann-Hahn condition.<sup>20,125</sup> Swept-frequency two-pulse phase modulation ( $\text{SW}_f\text{-TPPM}$ )<sup>24</sup> or small phase incremental alternation with 64 steps (SPINAL-64)<sup>25</sup> were applied for proton decoupling. The States-TPPI method<sup>126</sup> was used for phase-sensitive detection in all indirect dimensions.

During  $^{13}\text{C}$ - $^{13}\text{C}$  DARR mixing and the z-filter delay matched to multiples of the rotor period ( $\sim 200 \mu\text{s}$ ) in the  $^{15}\text{N}$ - $^{13}\text{C}$  ZF TEDOR experiment, protons were irradiated at an RF field matching the spinning frequency of 14.5 kHz. Table S3.2 gives a complete overview of the acquisition parameters per experiment and sample.

## Chapter 3 Co-sedimentation as the key to the investigation of wild-type FAT10 by MAS NMR spectroscopy

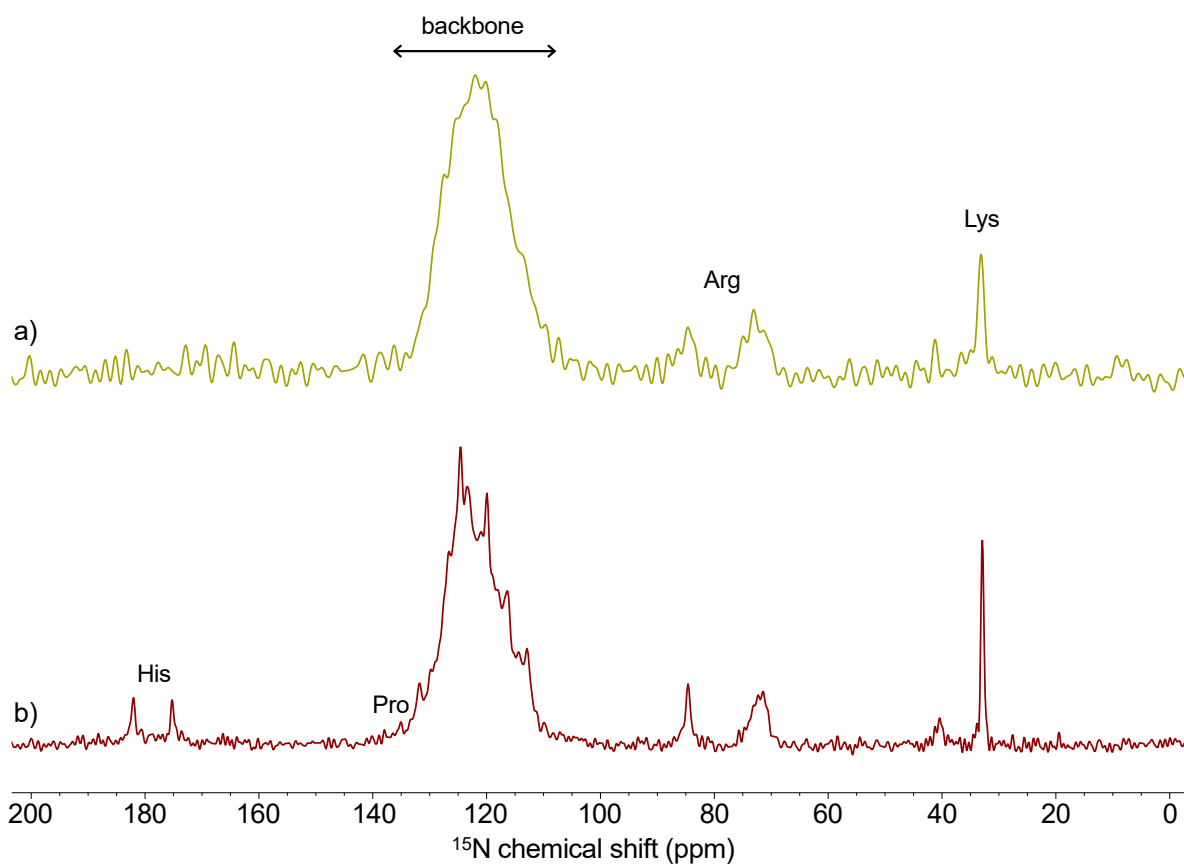
**Table S3.2:** Acquisition parameters for MAS NMR experiments.

Experiment	CP contact time (ms)		Radio frequency field strength (kHz)							Dwell time ( $\mu$ s)			Acquisition time (ms)			Carrier position (ppm)			Recycle delay (s)	
	$^1\text{H}$ - $^{13}\text{C}$ / $^1\text{H}$ - $^{15}\text{N}$ CP	$^{15}\text{N}$ - $^{13}\text{C}$ DCP	$^1\text{H}$ - $^{13}\text{C}$ / $^1\text{H}$ - $^{15}\text{N}$ CP		$^{15}\text{N}$ - $^{13}\text{C}$ DCP		$\pi/2$ - and $\pi$ -pulses		Decoupling											
			$^1\text{H}$ (linear) <sup>#</sup>	$^{13}\text{C}/^{15}\text{N}$	$^{15}\text{N}$	$^{13}\text{C}$ (tangent) <sup>#</sup>	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$	$^1\text{H}$	$^{13}\text{C}$	$^{13}\text{C}$	$^{15}\text{N}$	$^{13}\text{C}$	$^{13}\text{C}$	$^{15}\text{N}$	$^{13}\text{C}$	$^{13}\text{C}$		$^{15}\text{N}$
<b>U-<math>^{13}\text{C}</math>,<math>^{15}\text{N}</math>-FAT10-WT (amino acids 2-86)</b>																				
hC	1.2		71.1	50.0			83.3			71.4, SPINAL-64	7.00			14.3			100			2.5
hN	1.0		67.4	50.0			83.3			71.4, SPINAL-64			12.0			12.3			104	3.5
DARR	1.2		71.1	50.0			83.3	50.0		71.4, SPINAL-64	11.6	23.0		11.9	4.8		100	100		2.0
<b><math>^{13}\text{C}</math>,<math>^{15}\text{N}</math>-FAT10-WT (amino acids 5-86) in complex with natural abundance NUB1L</b>																				
hC	1.0		67.8	50.0			83.3			71.4, SPINAL-64	6.1			25.0			100			3.0
hN	0.9		59.1	41.7			83.3			71.4, SPINAL-64			12.0			24.6			104	3.0
DARR	1.0		67.8	50.0			83.3	50.0		71.4, SPINAL-64	11.3	23.0		23.1	11.4		100	100		2.0
ZF TEDOR	1.0		67.8	50.0			83.3	50.0	27.8	83.3/71.4, SWf-TPPM	12.0		69.0	24.6		7.6	100		104	2.5
NCO	0.75	1.25	68.2 (tangent)	41.7	6.95	6.05	85.9			90.0, SWf-TPPM/ CW	11.0		333.3	20.0		8.0	110 <sup>†</sup>		122	2.0
NCA	0.75	1.75	68.8 (tangent)	41.7	3.66	20.8	85.9			90.0, SWf-TPPM/ CW	11.0		333.3	20.0		8.0	110 <sup>†</sup>		122	2.0
NCOCX	0.75	1.25	68.2 (tangent)	41.7	6.95	6.05	85.9	60.6	41.7	90.0, SWf-TPPM/ CW	9.5	333.3	333.3	19.5	8.0	8.0	110	110 <sup>†</sup>	122	1.8
NCACX	0.75	2.0	72.3 (tangent)	41.7	3.66	21.4	89.3	60.6	41.7	95.0, SWf-TPPM/ CW	9.5	333.3	333.3	19.5	8.0	8.0	110	110 <sup>†</sup>	122	1.8

<sup>#</sup>Corresponds to 100 % of the ramp. <sup>†</sup>For  $^{15}\text{N}$ - $^{13}\text{C}$  SPECIFIC-CP, the CO carrier was centred at 172 and 174 ppm for NCO and NCOCX experiments, respectively. The CA carrier was centred at 55 and 56 ppm for NCA and NCACX experiments, respectively.

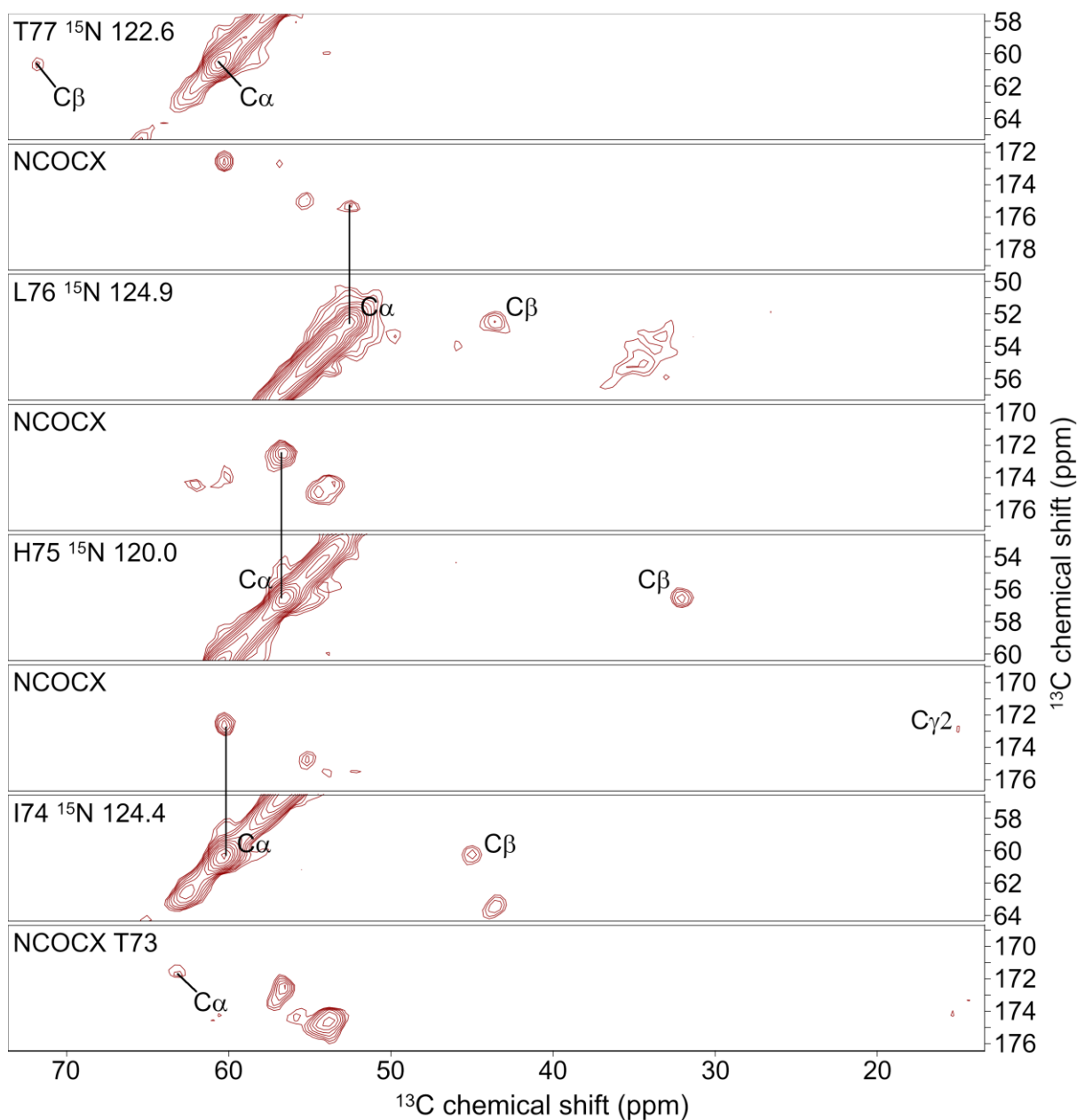
For free induction decays with unnecessary long acquisition times, the number of effective points was adjusted accordingly. 1D spectra were zero-filled and baseline corrected, but no apodization was applied. For 2D and 3D spectra, 60°- or 72°-shifted squared-sine bell apodization (corresponding to an offset of 0.33 or 0.40 in NMRPipe) and zero filling were used in all dimensions.

3.7.4  $^1\text{H}$ - $^{15}\text{N}$  cross-polarization spectra



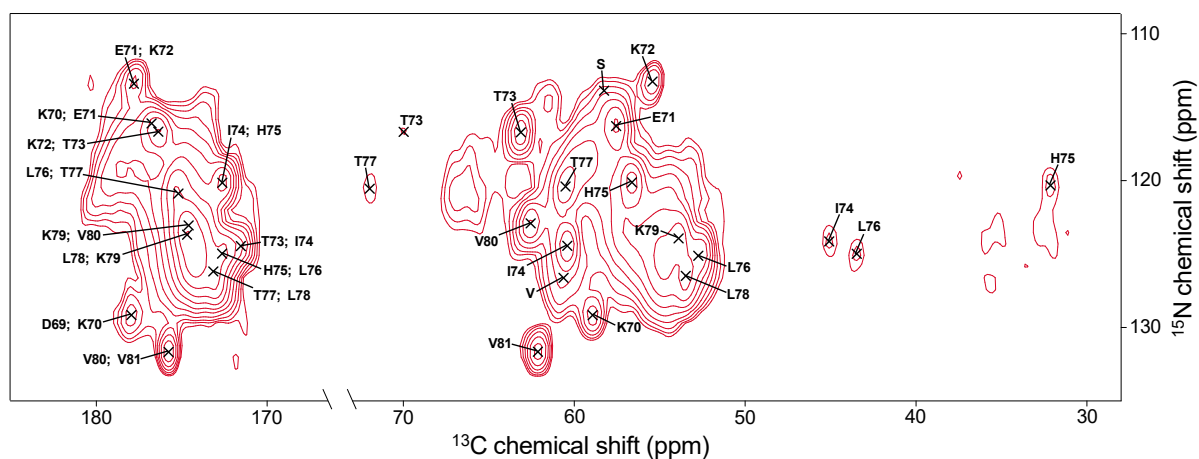
**Figure S3.2:**  $^1\text{H}$ - $^{15}\text{N}$  cross-polarization spectra of (a) N-FAT10-WT (amino acids 2-86) and (b) N-FAT10-WT (amino acids 5-86) in complex with NUB1L.

3.7.5 Strip plot of N-FAT10-WT in complex with NUB1L



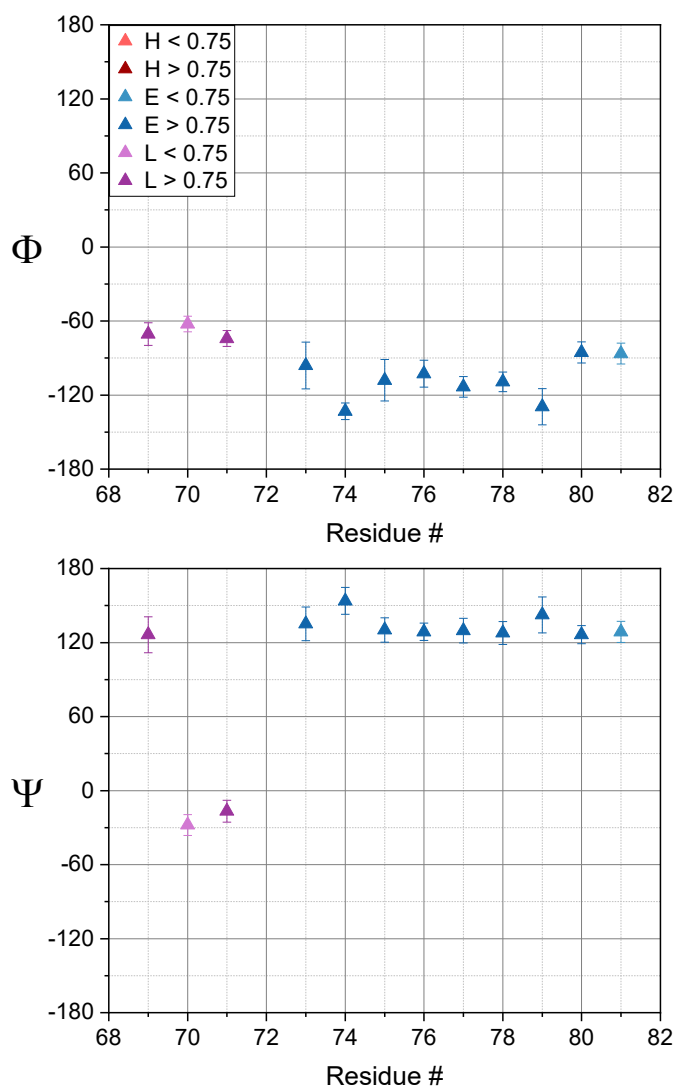
**Figure S3.3:**  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  NCOCX/NCACX representative strip plot for residues T73-T77 of N-FAT10-WT (amino acids 5-86) in complex with NUB1L.

3.7.6  $^{15}\text{N}$ - $^{13}\text{C}$  spectra of N-FAT10-WT in complex with NUB1L



**Figure S3.4:**  $^{15}\text{N}$ - $^{13}\text{C}$  NCO and NCA spectra of N-FAT10-WT (amino acids 5-86) in complex with NUB1L. Peak symbols of site-specifically or residue type assigned resonances are included and labelled accordingly.

3.7.7 Torsion angle analysis of N-FAT10-WT in complex with NUB1L



**Figure S3.5:** Backbone torsion angle and secondary structure prediction based on MAS NMR of N-FAT10-WT (amino acids 5-86) co-sedimented with NUB1L. Error bars correspond to the standard deviations of the  $\Phi$  and  $\Psi$  angles of the best matches in the TALOS-N database. For residue K72, there is no consensus among the database matches and, hence, no predicted torsion angles are shown. Secondary structure predictions are based on observed chemical shifts and classified as helix (H), strand (E) or loop (L). Predictions with a confidence  $> 0.75$  are highly reliable. Based on sequence information, the secondary structure of I68 is classified as loop. Based on observed chemical shifts, the secondary structure of K72 is classified as loop.

---

## References

1. Buxbaum, E. *Fundamentals of protein structure and function*, (Springer, 2007).
2. Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **7**, 95-99 (1963).
3. Reuter, J.A., Spacek, D.V. & Snyder, M.P. High-Throughput Sequencing Technologies. *Molecular Cell* **58**, 586-597 (2015).
4. Hensen, U. et al. Exploring Protein Dynamics Space: The Dynasome as the Missing Link between Protein Structure and Function. *PLOS ONE* **7**, e33931 (2012).
5. Blundell, T.L. The first resolution revolution in protein structure analysis: X-ray diffraction of polypeptide conformations and globular protein folds in 1950s and 1960s. *Progress in Biophysics and Molecular Biology* **167**, 32-40 (2021).
6. Cachau, R.E., Zhu, J. & Nicklaus, M.C. The upcoming subatomic resolution revolution. *Current Opinion in Structural Biology* **58**, 53-58 (2019).
7. Rathore, I., Mishra, V. & Bhaumik, P. Advancements in macromolecular crystallography: from past to present. *Emerging Topics in Life Sciences* **5**, 127-149 (2021).
8. Jumper, J. et al. Applying and improving AlphaFold at CASP14. *Proteins: Structure, Function, and Bioinformatics* **89**, 1711-1721 (2021).
9. Bai, X.-C., McMullan, G. & Scheres, S.H.W. How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences* **40**, 49-57 (2015).
10. Ahlawat, S., Mote, K.R., Lakomek, N.-A. & Agarwal, V. Solid-State NMR: Methods for Biological Solids. *Chemical Reviews* **122**, 9643-9737 (2022).
11. Reif, B., Ashbrook, S.E., Emsley, L. & Hong, M. Solid-state NMR spectroscopy. *Nature Reviews Methods Primers* **1**, 2 (2021).
12. Rabi, I.I., Zacharias, J.R., Millman, S. & Kusch, P. A New Method of Measuring Nuclear Magnetic Moment. *Physical Review* **53**, 318 (1938).
13. Andrew, E.R. & Szczesniak, E. A historical account of NMR in the solid state. *Progress in Nuclear Magnetic Resonance Spectroscopy* **28**, 11-36 (1995).
14. Levitt, M.H. *Spin dynamics: basics of nuclear magnetic resonance*, (John Wiley & Sons, 2008).
15. Andrew, E.R., Bradbury, A. & Eades, R.G. Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation. *Nature* **183**, 1802-1803 (1959).
16. Lowe, I.J. Free Induction Decays of Rotating Solids. *Physical Review Letters* **2**, 285-287 (1959).
17. Higman, V.A. Solid-state MAS NMR resonance assignment methods for proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy* **106-107**, 37-65 (2018).
18. Bertini, I., Luchinat, C., Parigi, G. & Ravera, E. SedNMR: On the Edge between Solution and Solid-State NMR. *Accounts of Chemical Research* **46**, 2059-2069 (2013).
19. Stringer, J.A. et al. Reduction of RF-induced sample heating with a scroll coil resonator structure for solid-state NMR probes. *Journal of Magnetic Resonance* **173**, 40-48 (2005).
20. Hartmann, S.R. & Hahn, E.L. Nuclear Double Resonance in the Rotating Frame. *Physical Review* **128**, 2042-2053 (1962).
21. Li, Y., Zhang, S., Wu, Z., Peng, X. & Fu, R. On the use of cross polarization in solid-state NMR: <sup>1</sup>H spin-lock versus adiabatic demagnetization in the rotating frame. *Magnetic Resonance Letters* **2**, 147-158 (2022).
22. Bennett, A.E., Rienstra, C.M., Auger, M., Lakshmi, K.V. & Griffin, R.G. Heteronuclear decoupling in rotating solids. *The Journal of Chemical Physics* **103**, 6951-6958 (1995).

## References

---

23. Detken, A., Hardy, E.H., Ernst, M. & Meier, B.H. Simple and efficient decoupling in magic-angle spinning solid-state NMR: the XiX scheme. *Chemical Physics Letters* **356**, 298-304 (2002).
24. Thakur, R.S., Kurur, N.D. & Madhu, P.K. Swept-frequency two-pulse phase modulation for heteronuclear dipolar decoupling in solid-state NMR. *Chemical Physics Letters* **426**, 459-463 (2006).
25. Fung, B.M., Khitrin, A.K. & Ermolaev, K. An Improved Broadband Decoupling Sequence for Liquid Crystals and Solids. *Journal of Magnetic Resonance* **142**, 97-101 (2000).
26. Paëpe, G.D. Dipolar Recoupling in Magic Angle Spinning Solid-State Nuclear Magnetic Resonance. *Annual Review of Physical Chemistry* **63**, 661-684 (2012).
27. Gullion, T. & Schaefer, J. Rotational-echo double-resonance NMR. *Journal of Magnetic Resonance (1969)* **81**, 196-200 (1989).
28. Schaefer, J. "Development of REDOR rotational-echo double-resonance NMR" by Terry Gullion and Jacob Schaefer [J. Magn. Reson. 81 (1989) 196–200]. *Journal of Magnetic Resonance* **213**, 421-422 (2011).
29. Jaroniec, C.P., Tounge, B.A., Herzfeld, J. & Griffin, R.G. Frequency Selective Heteronuclear Dipolar Recoupling in Rotating Solids: Accurate  $^{13}\text{C}$ - $^{15}\text{N}$  Distance Measurements in Uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled Peptides. *Journal of the American Chemical Society* **123**, 3507-3519 (2001).
30. Hing, A.W., Vega, S. & Schaefer, J. Transferred-echo double-resonance NMR. *Journal of Magnetic Resonance (1969)* **96**, 205-209 (1992).
31. Hing, A.W., Vega, S. & Schaefer, J. Measurement of Heteronuclear Dipolar Coupling by Transferred-Echo Double-Resonance NMR. *Journal of Magnetic Resonance, Series A* **103**, 151-162 (1993).
32. Jaroniec, C.P., Filip, C. & Griffin, R.G. 3D TEDOR NMR Experiments for the Simultaneous Measurement of Multiple Carbon-Nitrogen Distances in Uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -Labeled Solids. *Journal of the American Chemical Society* **124**, 10728-10742 (2002).
33. Szeverenyi, N.M., Sullivan, M.J. & Maciel, G.E. Observation of spin exchange by two-dimensional fourier transform  $^{13}\text{C}$  cross polarization-magic-angle spinning. *Journal of Magnetic Resonance (1969)* **47**, 462-475 (1982).
34. Takegoshi, K., Nakamura, S. & Terao, T.  $^{13}\text{C}$ - $^1\text{H}$  dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters* **344**, 631-637 (2001).
35. Takegoshi, K., Nakamura, S. & Terao, T.  $^{13}\text{C}$ - $^1\text{H}$  dipolar-driven  $^{13}\text{C}$ - $^{13}\text{C}$  recoupling without  $^{13}\text{C}$  rf irradiation in nuclear magnetic resonance of rotating solids. *The Journal of Chemical Physics* **118**, 2325-2341 (2003).
36. Bennett, A.E., Griffin, R.G., Ok, J.H. & Vega, S. Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *The Journal of Chemical Physics* **96**, 8624-8627 (1992).
37. Bayro, M.J. et al. Dipolar truncation in magic-angle spinning NMR recoupling experiments. *The Journal of Chemical Physics* **130**, 114506 (2009).
38. Castellani, F. et al. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 99-102 (2002).
39. Le Marchand, T. et al.  $^1\text{H}$ -Detected Biomolecular NMR under Fast Magic-Angle Spinning. *Chemical Reviews* **122**, 9943-10018 (2022).
40. Sun, Z. et al. Pushing the Boundaries of Resolution in Solid-State Nuclear Magnetic Resonance of Biomolecules with 160 kHz Magic-Angle Spinning. *Journal of the American Chemical Society* **147**, 19433-19437 (2025).
41. Showell, H.J. et al. The structure-activity relations of synthetic peptides as chemotactic factors and inducers of lysosomal secretion for neutrophils. *Journal of Experimental Medicine* **143**, 1154-1169 (1976).
42. Marasco, W.A. et al. Purification and identification of formyl-methionyl-leucyl-phenylalanine as the major peptide neutrophil chemotactic factor produced by *Escherichia coli*. *Journal of Biological Chemistry* **259**, 5430-9 (1984).

## References

---

43. Zech, S.G., Wand, A.J. & McDermott, A.E. Protein Structure Determination by High-Resolution Solid-State NMR Spectroscopy: Application to Microcrystalline Ubiquitin. *Journal of the American Chemical Society* **127**, 8618-8626 (2005).
44. Baldus, M., Petkova, A.T., Herzfeld, J. & Griffin, R.G. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Molecular Physics* **95**, 1197-1207 (1998).
45. Castellani, F., van Rossum, B.-J., Diehl, A., Rehbein, K. & Oschkinat, H. Determination of Solid-State NMR Structures of Proteins by Means of Three-Dimensional <sup>15</sup>N-<sup>13</sup>C-<sup>13</sup>C Dipolar Correlation Spectroscopy and Chemical Shift Analysis. *Biochemistry* **42**, 11476-11483 (2003).
46. Daviso, E., Eddy, M.T., Andreas, L.B., Griffin, R.G. & Herzfeld, J. Efficient resonance assignment of proteins in MAS NMR by simultaneous intra- and inter-residue 3D correlation spectroscopy. *Journal of Biomolecular NMR* **55**, 257-265 (2013).
47. Andreas, L.B., Eddy, M.T., Chou, J.J. & Griffin, R.G. Magic-Angle-Spinning NMR of the Drug Resistant S31N M2 Proton Transporter from Influenza A. *Journal of the American Chemical Society* **134**, 7215-7218 (2012).
48. Balch, W.E., Morimoto, R.I., Dillin, A. & Kelly, J.W. Adapting Proteostasis for Disease Intervention. *Science* **319**, 916-919 (2008).
49. Arkinson, C., Dong, K.C., Gee, C.L. & Martin, A. Mechanisms and regulation of substrate degradation by the 26S proteasome. *Nature Reviews Molecular Cell Biology* **26**, 104-122 (2025).
50. Goldberg, K. et al. Cell-autonomous innate immunity by proteasome-derived defence peptides. *Nature* **639**, 1032-1041 (2025).
51. Goldstein, G. et al. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences* **72**, 11-15 (1975).
52. Ciechanover, A., Hod, Y. & Hershko, A. A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communications* **81**, 1100-1105 (1978).
53. Wilkinson, K.D. & Audhya, T.K. Stimulation of ATP-dependent proteolysis requires ubiquitin with the COOH-terminal sequence Arg-Gly-Gly. *Journal of Biological Chemistry* **256**, 9235-9241 (1981).
54. Vijay-Kumar, S., Bugg, C.E. & Cook, W.J. Structure of ubiquitin refined at 1.8 Å resolution. *Journal of Molecular Biology* **194**, 531-544 (1987).
55. Aichem, A. & Groettrup, M. The ubiquitin-like modifier FAT10 – much more than a proteasome-targeting signal. *Journal of Cell Science* **133**, jcs246041 (2020).
56. Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of Proteins by Ubiquitin and Ubiquitin-Like Proteins. *Annual Review of Cell and Developmental Biology* **22**, 159-180 (2006).
57. Hershko, A. & Ciechanover, A. THE UBIQUITIN SYSTEM FOR PROTEIN DEGRADATION. *Annual Review of Biochemistry* **61**, 761-807 (1992).
58. Thrower, J.S., Hoffman, L., Rechsteiner, M. & Pickart, C.M. Recognition of the polyubiquitin proteolytic signal. *The EMBO Journal* **19**, 94-102 (2000).
59. Reyes-Turcu, F.E., Ventii, K.H. & Wilkinson, K.D. Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annual review of biochemistry* **78**, 363-397 (2009).
60. Jariel-Encontre, I., Bossis, G. & Piechaczyk, M. Ubiquitin-independent degradation of proteins by the proteasome. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1786**, 153-177 (2008).
61. Eroles, J. & Coffino, P. Ubiquitin-independent proteasomal degradation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**, 216-221 (2014).
62. Gu, X. et al. The midnolin-proteasome pathway catches proteins for ubiquitination-independent degradation. *Science* **381**, eadh5021 (2023).
63. Nardone, C. et al. Structural basis for the midnolin-proteasome pathway and its role in suppressing myeloma. *Molecular Cell* (2025).

## References

---

64. Fan, W., Cai, W., Parimoo, S., Lennon, G.G. & Weissman, S.M. Identification of seven new human MHC class I region genes around the HLA-F locus. *Immunogenetics* **44**, 97-103 (1996).
65. Chiu, Y.-H., Sun, Q. & Chen, Z.J. E1-L2 Activates Both Ubiquitin and FAT10. *Molecular Cell* **27**, 1014-1023 (2007).
66. Groettrup, M., Pelzer, C., Schmidtke, G. & Hofmann, K. Activating the ubiquitin family: UBA6 challenges the field. *Trends in Biochemical Sciences* **33**, 230-237 (2008).
67. Aichem, A. et al. USE1 is a bispecific conjugating enzyme for ubiquitin and FAT10, which FAT10ylates itself in cis. *Nature Communications* **1**, 13 (2010).
68. Roverato, N.D. et al. Parkin is an E3 ligase for the ubiquitin-like modifier FAT10, which inhibits Parkin activation and mitophagy. *Cell Reports* **34**, 108857 (2021).
69. Hipp, M.S., Kalveram, B., Raasi, S., Groettrup, M. & Schmidtke, G. FAT10, a ubiquitin-independent signal for proteasomal degradation. *Molecular and cellular biology* **25**, 3483-3491 (2005).
70. Aichem, A. et al. The structure of the ubiquitin-like modifier FAT10 reveals an alternative targeting mechanism for proteasomal degradation. *Nature Communications* **9**, 3321 (2018).
71. Arkinson, C. et al. NUB1 traps unfolded FAT10 for ubiquitin-independent degradation by the 26S proteasome. *Nature Structural & Molecular Biology* **32**, 1752-1765 (2025).
72. Hipp, M.S., Raasi, S., Groettrup, M. & Schmidtke, G. NEDD8 Ultimate Buster-1L Interacts with the Ubiquitin-like Protein FAT10 and Accelerates Its Degradation. *Journal of Biological Chemistry* **279**, 16503-16510 (2004).
73. Rani, N., Aichem, A., Schmidtke, G., Kreft, S.G. & Groettrup, M. FAT10 and NUB1L bind to the VWA domain of Rpn10 and Rpn1 to enable proteasome-mediated proteolysis. *Nature Communications* **3**, 749 (2012).
74. Schmidtke, G. et al. The UBA Domains of NUB1L Are Required for Binding but Not for Accelerated Degradation of the Ubiquitin-like Modifier FAT10. *Journal of Biological Chemistry* **281**, 20045-20054 (2006).
75. Theng, S.S. et al. Disruption of FAT10–MAD2 binding inhibits tumor progression. *Proceedings of the National Academy of Sciences* **111**, E5282-E5291 (2014).
76. Bates, E.E.M. et al. Identification and analysis of a novel member of the ubiquitin family expressed in dendritic cells and mature B cells. *European Journal of Immunology* **27**, 2471-2477 (1997).
77. Lukasiak, S. et al. Proinflammatory cytokines cause FAT10 upregulation in cancers of liver and colon. *Oncogene* **27**, 6068-6074 (2008).
78. Schregle, R. et al. The expression profile of the ubiquitin-like modifier FAT10 in immune cells suggests cell type-specific functions. *Immunogenetics* **70**, 429-438 (2018).
79. Liu, Y.C. et al. A MHC-encoded ubiquitin-like protein (FAT10) binds noncovalently to the spindle assembly checkpoint protein MAD2. *Proceedings of the National Academy of Sciences* **96**, 4313-4318 (1999).
80. Raasi, S., Schmidtke, G., de Giuli, R. & Groettrup, M. A ubiquitin-like protein which is synergistically inducible by interferon- $\gamma$  and tumor necrosis factor- $\alpha$ . *European Journal of Immunology* **29**, 4030-6 (1999).
81. Ebstein, F., Lehmann, A. & Kloetzel, P.-M. The FAT10- and ubiquitin-dependent degradation machineries exhibit common and distinct requirements for MHC class I antigen presentation. *Cellular and molecular life sciences : CMLS* **69**, 2443-2454 (2012).
82. Aichem, A. et al. The proteomic analysis of endogenous FAT10 substrates identifies p62/SQSTM1 as a substrate of FAT10ylation. *Journal of Cell Science* **125**, 4576-4585 (2012).
83. Lee, C.G.L. et al. Expression of the FAT10 gene is highly upregulated in hepatocellular carcinoma and other gastrointestinal and gynecological cancers. *Oncogene* **22**, 2592-2603 (2003).

## References

---

84. Schnell, L., Zubrod, A., Catone, N., Bialas, J. & Aichele, A. Tumor necrosis factor mediates UBE1-independent FAT10ylation under inflammatory conditions. *Life Science Alliance* **6**, e202301985 (2023).
85. Schmidtke, G., Kalveram, B. & Groettrup, M. Degradation of FAT10 by the 26S proteasome is independent of ubiquitylation but relies on NUB1L. *FEBS Letters* **583**, 591-594 (2009).
86. Brockmann, F. et al. FAT10 and NUB1L cooperate to activate the 26S proteasome. *Life Science Alliance* **6**, e202201463 (2023).
87. Buchsbaum, S., Bercovich, B., Ziv, T. & Ciechanover, A. Modification of the inflammatory mediator LRRFIP2 by the ubiquitin-like protein FAT10 inhibits its activity during cellular response to LPS. *Biochemical and Biophysical Research Communications* **428**, 11-16 (2012).
88. Nguyen, N.T.H., Now, H., Kim, W.-J., Kim, N. & Yoo, J.-Y. Ubiquitin-like modifier FAT10 attenuates RIG-I mediated antiviral signaling by segregating activated RIG-I from its signaling platform. *Scientific Reports* **6**, 23377 (2016).
89. Guo, C. et al. Structural basis of protein condensation on microtubules underlying branching microtubule nucleation. *Nature Communications* **14**, 3682 (2023).
90. El Mammeri, N., Duan, P. & Hong, M. Pseudo-Phosphorylated Tau Forms Paired Helical Filaments in the Presence of High-Curvature Cholesterol-Containing Lipid Membranes. *Journal of the American Chemical Society* **147**, 2510-2522 (2025).
91. Barbet-Massin, E., Huang, C.-T., Daebel, V., Hsu, S.-T.D. & Reif, B. Site-Specific Solid-State NMR Studies of "Trigger Factor" in Complex with the Large Ribosomal Subunit 50S. *Angewandte Chemie International Edition* **54**, 4367-4369 (2015).
92. Gardienet, C. et al. Solid-state NMR chemical-shift perturbations indicate domain reorientation of the DnaG primase in the primosome of *Helicobacter pylori*. *Journal of Biomolecular NMR* **64**, 189-195 (2016).
93. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.10.04.463034 (2022).
94. Mandal, A., Boatz, J.C., Wheeler, T.B. & van der Wel, P.C.A. On the use of ultracentrifugal devices for routine sample preparation in biomolecular magic-angle-spinning NMR. *Journal of Biomolecular NMR* **67**, 165-178 (2017).
95. Hou, G., Yan, S., Trébosc, J., Amoureux, J.-P. & Polenova, T. Broadband homonuclear correlation spectroscopy driven by combined R2nv sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *Journal of Magnetic Resonance* **232**, 18-30 (2013).
96. Delaglio, F. et al. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR* **6**, 277-293 (1995).
97. Skinner, S.P. et al. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *Journal of Biomolecular NMR* **66**, 111-124 (2016).
98. Wishart, D.S., Sykes, B.D. & Richards, F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of Molecular Biology* **222**, 311-333 (1991).
99. Spera, S. & Bax, A. Empirical correlation between protein backbone conformation and C $\alpha$  and C $\beta$   $^{13}\text{C}$  nuclear magnetic resonance chemical shifts. *Journal of the American Chemical Society* **113**, 5490-5492 (1991).
100. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *Journal of Biomolecular NMR* **56**, 227-241 (2013).
101. Meng, E.C. et al. UCSF ChimeraX: Tools for structure building and analysis. *Protein Science* **32**, e4792 (2023).
102. Willard, L. et al. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Research* **31**, 3316-3319 (2003).
103. Martin, R.W. & Zilm, K.W. Preparation of protein nanocrystals and their characterization by solid state NMR. *Journal of Magnetic Resonance* **165**, 162-174 (2003).

104. Kraus, J., Sarkar, S., Quinn, C.M. & Polenova, T. Chapter Two - Solid-state NMR spectroscopy of microcrystalline proteins. in *Annual Reports on NMR Spectroscopy*, Vol. 102 (ed. Webb, G.A.) 81-151 (Academic Press, 2021).
105. Frericks Schmidt, H.L., Shah, G.J., Sperling, L.J. & Rienstra, C.M. NMR Determination of Protein  $pK_a$  Values in the Solid State. *The Journal of Physical Chemistry Letters* **1**, 1623-1628 (2010).
106. Platzner, G., Okon, M. & McIntosh, L.P. pH-dependent random coil  $^1H$ ,  $^{13}C$ , and  $^{15}N$  chemical shifts of the ionizable amino acids: a guide for protein  $pK_a$  measurements. *Journal of Biomolecular NMR* **60**, 109-129 (2014).
107. Morris, G.A. & Freeman, R. Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* **101**, 760-762 (1979).
108. Truongvan, N., Li, S., Misra, M., Kuhn, M. & Schindelin, H. Structures of UBA6 explain its dual specificity for ubiquitin and FAT10. *Nature Communications* **13**, 4789 (2022).
109. Cao, J., Aichem, A., Basler, M., Alvarez Salinas, G.O. & Schmidtke, G. Phosphorylated FAT10 Is More Efficiently Conjugated to Substrates, Does Not Bind to NUB1L, and Does Not Alter Degradation by the Proteasome. *Biomedicines* **12**, 2795 (2024).
110. Remaut, H. & Waksman, G. Protein-protein interaction through  $\beta$ -strand addition. *Trends in Biochemical Sciences* **31**, 436-444 (2006).
111. Le Marchand, T. et al. Conformational dynamics in crystals reveal the molecular bases for D76N beta-2 microglobulin aggregation propensity. *Nature Communications* **9**, 1658 (2018).
112. Gauto, D.F. et al. Functional control of a 0.5 MDa TET aminopeptidase by a flexible loop revealed by MAS NMR. *Nature Communications* **13**, 1927 (2022).
113. Troussicot, L., Vallet, A., Molin, M., Burmann, B.M. & Schanda, P. Disulfide-Bond-Induced Structural Frustration and Dynamic Disorder in a Peroxiredoxin from MAS NMR. *Journal of the American Chemical Society* **145**, 10700-10711 (2023).
114. Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends in Biochemical Sciences* **33**, 2-8 (2008).
115. Alderson, T.R., Pritišanac, I., Kolarić, Đ., Moses, A.M. & Forman-Kay, J.D. Systematic identification of conditionally folded intrinsically disordered regions by AlphaFold2. *Proceedings of the National Academy of Sciences* **120**, e2304302120 (2023).
116. Bah, A. et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106-109 (2015).
117. Fukuyo, A., In, Y., Ishida, T. & Tomoo, K. Structural scaffold for eIF4E binding selectivity of 4E-BP isoforms: crystal structure of eIF4E binding region of 4E-BP2 and its comparison with that of 4E-BP1. *Journal of Peptide Science* **17**, 650-657 (2011).
118. Sekiyama, N. et al. Molecular mechanism of the dual activity of 4EGI-1: Dissociating eIF4G from eIF4E but stabilizing the binding of unphosphorylated 4E-BP1. *Proceedings of the National Academy of Sciences* **112**, E4036-E4045 (2015).
119. Guo, C., Hou, G., Lu, X. & Polenova, T. Mapping protein-protein interactions by double-REDOR-filtered magic angle spinning NMR spectroscopy. *Journal of Biomolecular NMR* **67**, 95-108 (2017).
120. Rovó, P. Recent advances in solid-state relaxation dispersion techniques. *Solid State Nuclear Magnetic Resonance* **108**, 101665 (2020).
121. Thurber, K.R. & Tycko, R. Measurement of sample temperatures under magic-angle spinning from the chemical shift and spin-lattice relaxation rate of  $^{79}Br$  in KBr powder. *Journal of Magnetic Resonance* **196**, 84-87 (2009).
122. Morcombe, C.R. & Zilm, K.W. Chemical shift referencing in MAS solid state NMR. *Journal of Magnetic Resonance* **162**, 479-486 (2003).
123. Harris, R.K. et al. Further Conventions for NMR Shielding and Chemical Shifts (IUPAC Recommendations 2008) Copyright IUPAC (2008). Reprinted with permission from Pure Appl. Chem., 2008, 80, 59-84, which can be freely downloaded from [www.iupac.org/publications/pac/80/1/0059](http://www.iupac.org/publications/pac/80/1/0059), provided IUPAC copyright is

- acknowledged in any use. Membership of the relevant IUPAC Committee is listed therein. in *eMagRes*.
124. Pines, A., Gibby, M.G. & Waugh, J.S. Proton-enhanced NMR of dilute spins in solids. *The Journal of Chemical Physics* **59**, 569-590 (1973).
  125. Stejskal, E.O., Schaefer, J. & Waugh, J.S. Magic-angle spinning and polarization transfer in proton-enhanced NMR. *Journal of Magnetic Resonance (1969)* **28**, 105-112 (1977).
  126. Marion, D., Ikura, M., Tschudin, R. & Bax, A. Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *Journal of Magnetic Resonance (1969)* **85**, 393-399 (1989).
  127. Farías-Rico, J.A. & Mourra-Díaz, C.M. A Short Tale of the Origin of Proteins and Ribosome Evolution. *Microorganisms* **10**, 2115 (2022).
  128. Morris, R., Black, K.A. & Stollar, E.J. Uncovering protein function: from classification to complexes. *Essays in Biochemistry* **66**, 255-285 (2022).
  129. Uversky, V.N. Introduction to Intrinsically Disordered Proteins (IDPs). *Chemical Reviews* **114**, 6557-6560 (2014).
  130. Trivedi, R. & Nagarajaram, H.A. Intrinsically Disordered Proteins: An Overview. *International Journal of Molecular Sciences* **23**, 14050 (2022).
  131. Dunker, A.K. et al. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* **19**, 26-59 (2001).
  132. Eliezer, D. Biophysical characterization of intrinsically disordered proteins. *Current Opinion in Structural Biology* **19**, 23-30 (2009).
  133. Negi, H., Ravichandran, A., Dasgupta, P., Reddy, S. & Das, R. Plasticity of the proteasome-targeting signal Fat10 enhances substrate degradation. *eLife* **13**, e91122 (2024).
  134. Weiss, C. et al. Intermolecular  $\beta$ -sheet formation guides the interaction between ubiquitin-like modifier FAT10 and adapter protein NUB1L. *bioRxiv*, 2025.07.31.667942 (2025).
  135. Thompson, L.K. Solid-state NMR studies of the structure and mechanisms of proteins. *Current Opinion in Structural Biology* **12**, 661-669 (2002).
  136. Opella, S.J. & Marassi, F.M. Structure Determination of Membrane Proteins by NMR Spectroscopy. *Chemical Reviews* **104**, 3587-3606 (2004).
  137. Middleton, D.A. NMR studies of amyloid interactions. *Progress in Nuclear Magnetic Resonance Spectroscopy* **144-145**, 63-96 (2024).
  138. Tycko, R. The evolving role of solid state nuclear magnetic resonance methods in studies of amyloid fibrils. *Current Opinion in Structural Biology* **92**, 103043 (2025).
  139. Xue, K. et al. Towards a native environment: structure and function of membrane proteins in lipid bilayers by NMR. *Chemical Science* **12**, 14332-14342 (2021).
  140. Saad, A. & Bechinger, B. Solid-state NMR spectroscopy for structural studies of polypeptides and lipids in extended physiological membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1866**, 184162 (2024).
  141. Porat-Dahlerbruch, G., Goldbourt, A. & Polenova, T. Virus Structures and Dynamics by Magic-Angle Spinning NMR. *Annual Review of Virology* **8**, 219-237 (2021).
  142. Lecoq, L. et al. Molecular elucidation of drug-induced abnormal assemblies of the hepatitis B virus capsid protein by solid-state NMR. *Nature Communications* **14**, 471 (2023).
  143. Bertini, I. et al. Solid-state NMR of proteins sedimented by ultracentrifugation. *Proceedings of the National Academy of Sciences* **108**, 10396-10399 (2011).
  144. Gardiennet, C. et al. A Sedimented Sample of a 59 kDa Dodecameric Helicase Yields High-Resolution Solid-State NMR Spectra. *Angewandte Chemie International Edition* **51**, 7855-7858 (2012).
  145. Barna, J.C.J., Laue, E.D., Mayger, M.R., Skilling, J. & Worrall, S.J.P. Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments. *Journal of Magnetic Resonance (1969)* **73**, 69-77 (1987).

## References

---

146. Robson, S., Arthanari, H., Hyberts, S.G. & Wagner, G. Chapter Ten - Nonuniform Sampling for NMR Spectroscopy. in *Methods in Enzymology*, Vol. 614 (ed. Wand, A.J.) 263-291 (Academic Press, 2019).
147. Stern, A.S., Donoho, D.L. & Hoch, J.C. NMR data processing using iterative thresholding and minimum l1-norm reconstruction. *Journal of Magnetic Resonance* **188**, 295-300 (2007).
148. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
149. Igumenova, T.I. et al. Assignments of Carbon NMR Resonances for Microcrystalline Ubiquitin. *Journal of the American Chemical Society* **126**, 6720-6727 (2004).
150. Fritzsche, K.J., Yang, Y., Schmidt-Rohr, K. & Hong, M. Practical use of chemical shift databases for protein solid-state NMR: 2D chemical shift maps and amino-acid assignment with secondary-structure information. *Journal of Biomolecular NMR* **56**, 155-167 (2013).
151. Liu, Z. et al. Noncovalent dimerization of ubiquitin. *Angewandte Chemie International Edition* **51**, 469-472 (2012).
152. Fragai, M., Luchinat, C., Parigi, G. & Ravera, E. Practical considerations over spectral quality in solid state NMR spectroscopy of soluble proteins. *Journal of Biomolecular NMR* **57**, 155-166 (2013).
153. *Biophysics of Molecular Chaperones: Function, Mechanisms and Client Protein Interactions*, 1-394 (Royal Soc Chemistry, CAMBRIDGE, 2024).
154. Prakash, S., Tian, L., Ratliff, K.S., Lehotzky, R.E. & Matouschek, A. An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nature Structural & Molecular Biology* **11**, 830-837 (2004).
155. Beskow, A. et al. A Conserved Unfoldase Activity for the p97 AAA-ATPase in Proteasomal Degradation. *Journal of Molecular Biology* **394**, 732-746 (2009).
156. Peth, A., Uchiki, T. & Goldberg, A.L. ATP-Dependent Steps in the Binding of Ubiquitin Conjugates to the 26S Proteasome that Commit to Degradation. *Molecular Cell* **40**, 671-681 (2010).
157. Osborn Popp, T.M. et al. 3D-Printable centrifugal devices for biomolecular solid state NMR rotors. *Journal of Magnetic Resonance* **354**, 107524 (2023).

## Curriculum Vitae

### Charlotte Weiss

Stadionstraße 16

78628 Rottweil, Germany

Date of birth: 12.06.1995

Place of birth: Rottweil, Germany

Phone: 0049173-9522156

Email: charlotte.weiss@uni-konstanz.de

LinkedIn: [www.linkedin.com/in/charlotte-weiß-76b9151b8](http://www.linkedin.com/in/charlotte-weiß-76b9151b8)

### Education and Research Experience

- 06/2020 – 06/2025 **Doctor of Natural Sciences (Dr. rer. nat.), Department of Chemistry, University of Konstanz, Konstanz, Germany**  
Doctoral thesis: *Magic-angle spinning NMR investigation of the interaction between FAT10 and NUB1L*  
Advisor: *Dr. Guinevere Mathies*
- Structural investigation of the N-domain of the protein FAT10 in its cysteine-free and wild-type forms as well as in complex with the protein NUB1L using magic-angle spinning NMR spectroscopy at high fields
- 10/2018 – 12/2020 **Master of Life Science, Department of Chemistry, University of Konstanz, Konstanz, Germany**  
Final grade: 1.0 | Konstanz Fast Track  
Master thesis: *Structural study of wild-type FAT10 by MAS NMR spectroscopy*
- Mitarbeiterpraktikum in the group of Dr. Guinevere Mathies: Carbon correlation spectroscopy of glutamine
  - Mitarbeiterpraktikum in the group of Prof. Dr. Stefan Mecking: Synthesis of PE-based vitrimers and investigation of their mechanical properties
  - Advanced Course Plant Ecology
  - Advanced Course Bioinformatics and X-Ray Structure Analysis
  - Internship in the School of Chemistry at University College Dublin
- 11/2018 – 12/2018 **Research Assistant, Hauser Group Biophysical Chemistry, University of Konstanz, Konstanz, Germany**
- Synthesis of alanine peptides
  - Spectroscopic analysis of synthesized peptides (CD and FTIR spectroscopy)
  - <sup>18</sup>O-labelling of glutamine
  - Mass spectrometry and purification *via* HPLC
- 10/2015 – 10/2018 **Bachelor of Life Science, Department of Chemistry, University of Konstanz, Konstanz, Germany**  
Final grade: 1.5  
Bachelor thesis: *Modellpeptide mit IR-spezifischen Sonden zur Untersuchung von Proteinfaltung*

### Teaching and Mentoring

10/2021 – 04/2022 **Tutor for Physical Chemistry II WS 21/22**

- Preparation and revision of problem sets
- Supervision of tutorials
- Exam drafting and marking

05/2021 – 08/2021 **Supervisor of Bachelor's thesis**

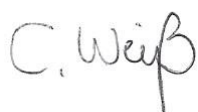
- Guidance of the student during his work towards the thesis (Title: MAS NMR Spektroskopie der N-Domäne von FAT10 bei 800 MHz)
- Teaching of protein expression/purification and operation of the 800 MHz spectrometer

01/2021 – 04/2021 **Tutor for Physical Chemistry II WS 20/21**

- Preparation and revision of problem sets
- Exam drafting and marking

### Conference Contributions

- 11/2023 SFB 969 Symposium 2023, Konstanz, Germany  
*New insights into structure, dynamics, and function of the ubiquitin-like modifier FAT10* (Poster)
- 09/2023 Young Investigators in Magnetic Resonances Workshop 2023, Konstanz, Germany  
*Probing the ubiquitin-like modifier FAT10 for non-covalent binding sites to the adaptor protein NUB1L by MAS NMR at 800 MHz* (Talk)
- 07/2023 19<sup>th</sup> European Magnetic Resonance Congress, Glasgow, Scotland, The United Kingdom  
*Probing the ubiquitin-like modifier FAT10 for non-covalent binding sites to the adaptor protein NUB1L by MAS NMR at 800 MHz* (Poster)
- 08/2022 Konstanz Research School Chemical Biology Retreat 2022, Bad Herrenalb, Germany  
*MAS NMR structural study of the FAT10 N-domain at 800 MHz* (Poster)
- 07/2022 18<sup>th</sup> European Magnetic Resonance Congress, Utrecht, The Netherlands  
*MAS NMR structural study of the FAT10 N-domain at 800 MHz* (Poster)
- 06/2022 Ampere NMR school 2022, Zakopane, Poland  
*MAS NMR structural study of the FAT10 N-domain at 800 MHz* (Poster)
- 04/2022 SFB 969 Symposium 2022, Konstanz, Germany  
*800 MHz MAS NMR structural study of the FAT10 N-domain* (Poster)
- 08/2021 Konstanz Research School Chemical Biology Retreat 2021, Gültstein, Germany  
*Structural study of wild-type FAT10 by MAS NMR spectroscopy* (Poster)
- 07/2021 17<sup>th</sup> European Magnetic Resonance Congress (Online)  
*Structural study of wild-type FAT10 by MAS NMR spectroscopy* (Poster)



Charlotte Weiss