

A combination of novel solid-state NMR methods and related software to study molecular assemblies and biomolecules

Sabine Gradmann

ISBN 978-90-5335-722-4

Doctoral thesis

A combination of novel solid-state NMR methods and related software to study molecular assemblies and biomolecules

Sabine Gradmann

NMR Spectroscopy Group, Bijvoet Center for Biomolecular Research

Utrecht University, The Netherlands

September 2013

**A combination of novel solid-state NMR methods and related
software
to study molecular assemblies and biomolecules**

Een combinatie van nieuwe vaste stof NMR methoden en gerelateerde software voor onderzoek naar moleculaire assemblages en biomoleculen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 18 september 2013 des middags te 12.45 uur

door

Sabine Hedwig Erika Gradmann

geboren op 1 maart 1983
te Bayreuth, Duitsland

Promotor: Prof. dr. M. Baldus

Ja, man muß seinen Traum finden, dann wird der Weg leicht. Aber es gibt keinen immerwährenden Traum, jeden löst ein neuer ab, und keinen darf man festhalten wollen.

- Hermann Hesse, Demian

Table of Contents

Chapter 1:	General Introduction	1
Chapter 2:	Rapid prediction of multi-dimensional NMR data sets	13
Chapter 3:	Solid-state NMR characterization of functional hydrogels formed by <i>Xenopus</i> Nup98 FG nucleoporins	31
Chapter 4:	Structural investigation of a homopolymeric protein hydrogel via solid-state NMR	51
Chapter 5:	Probing π - π stacking in thermosensitive polymeric micelles with a potential application as drug delivery systems	65
Chapter 6:	Investigation of Sensory Rhodopsin II and the NpSRII/NpHtrII complex by novel preparation and solid-state NMR methods	77
Chapter 7:	Summary and Outlook	99
Supporting Information		107
Acknowledgements		111
Publications		113
CV		115

CHAPTER 1

General Introduction

Solid-state NMR

Solid-state NMR (ssNMR) is a widely applicable method to infer structural and dynamical information in applications ranging from gels and powders to (proteo)liposomes and even cellular preparations [1, 2]. Typical spectral patterns result from the following main interactions for $\frac{1}{2}$ nuclei:

- The chemical shift is based on the interaction of the external magnetic field with the nucleus via its surrounding electrons. Due to their different electronic environment, the nuclei experience slightly diverse magnetic fields. In ssNMR, the chemical shift has an isotropic and orientation-dependent anisotropic (CSA) contribution.
- Dipolar couplings are magnetic interactions between nuclei through space. They depend on their distance and orientation.
- Scalar J-couplings are indirect magnetic interactions of nuclei through electron bonds.

The orientation-dependent anisotropic parts of the above mentioned interactions often lead to broad lines in the solid state.

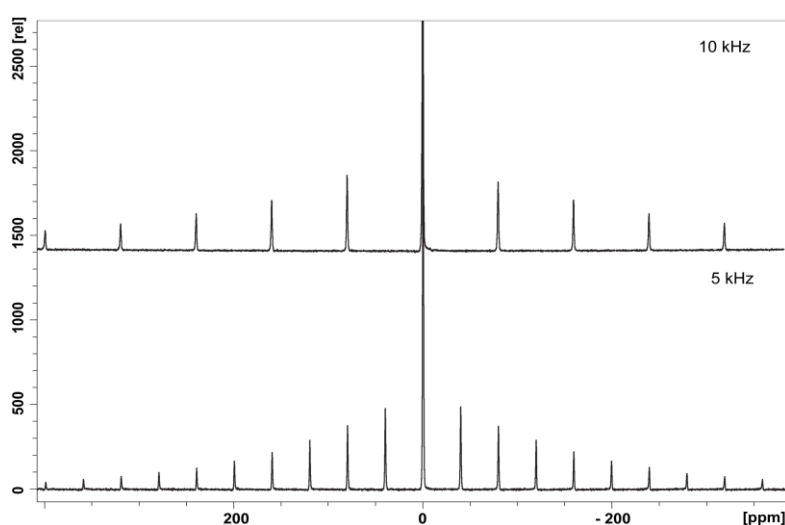


Figure 1: ^{79}Br NMR spectra of KBr at different spinning speeds.

Therefore, magic angle spinning (MAS, [3]) is applied to reduce these contributions by simulating molecular tumbling. The set angle value of 54.7° between the rotor (sample) axis and the external magnetic field is determined by the orientation dependency of the CSA and dipolar couplings that follow the second order Legendre polynomial ($3\cos^2\theta - 1 = 0$). Anisotropic contributions are averaged out more and more effectively with

increasing rotational speed. Remaining residual dipolar couplings and CSA interactions are present in the spectra as so-called spinning sidebands. An example is shown in Figure 1, where ^{79}Br spectra of KBr at 5 and 10 kHz spinning speed are displayed. Sidebands around the signal at the center occur in multiples of the spinning speed. As these interactions also provide valuable information (e.g. distance constraints), they are partially reintroduced via specific pulse sequence building blocks in the experiments.

Due to the different types of interactions, ssNMR allows to monitor a large range of motional regimes (from s to ns scale). Tailored experiments can map the dynamical behavior of proteins as well as exchange phenomena between different states or, in the context of membrane proteins, with the lipid environment. Beside the direct measurement of relaxation and correlation times, the nature of the magnetization transfer from protons to nuclei with a lower gyromagnetic ratio provides a further possibility to filter out specific motional regimes. Dipolar-based transfer (cross polarization, CP, [4]) and scalar-based transfer (insensitive nuclei enhanced by polarization transfer, INEPT) probes rigid and highly mobile parts of the sample, respectively [5].

Besides probing dynamical features ssNMR has gained importance on studying the structure of proteins in their native (-like) environment. The first step towards 3D structure determination usually includes the assignment of ^{13}C and ^{15}N resonances. Therefore, a set of standard 2D and 3D experiments are performed [6].

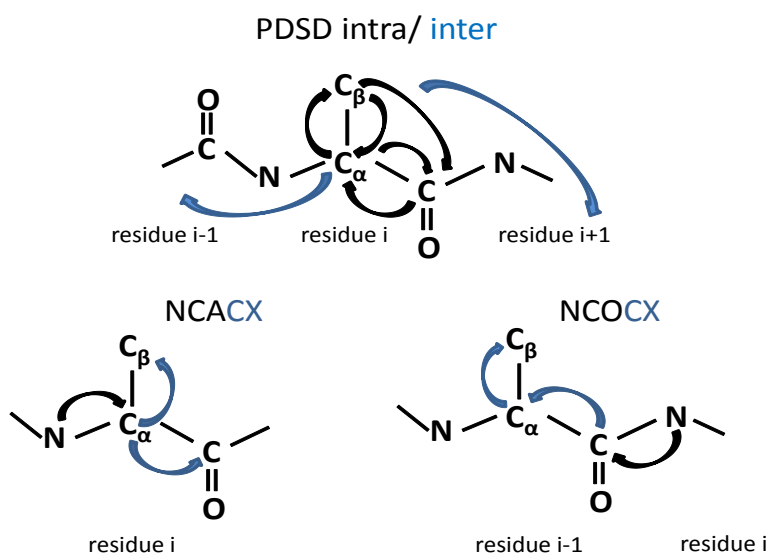


Figure 2: Schemes representing the magnetization transfer for standard 2D/ 3D ssNMR experiments which are commonly used for ^{13}C and ^{15}N assignments.

Homonuclear (^{13}C , ^{13}C) correlation experiments like Proton Driven Spin Diffusion (PDS, [7]) provide intra- and inter-residue information as shown via the black and blue arrows in Figure 2. Magnetization transfer takes place in both directions e.g. $\text{C}_\alpha \rightarrow \text{C}_\beta$ and $\text{C}_\beta \rightarrow \text{C}_\alpha$ which results in symmetric spectra. On the contrary, the magnetization is specifically transferred from $^{15}\text{N}_i$ to $\text{C}_{\alpha i}$ or CO_{i-1} in the NCACX or NCOCX experiment. Afterwards, it propagates further to other close-by carbons (blue arrows in Figure 2). In addition, distance information is obtained e.g. by CHHC/NHHC experiments [8], where magnetization is transferred from a carbon/nitrogen to another carbon via close-by protons through space. Applied on mixed labeled samples (a 1:1 mixture of uniformly ^{15}N and ^{13}C labeled monomers), NHHC experiments provide an insight into the interface and general assembly of proteins. The chemical shift assignments as well as the distance information are used as dihedral and distance restraints by structure calculation programs like CNS [9] or docking programs as HADDOCK [10] to generate protein structures.

Recently, combinations of new preparation methods e.g. perdeuteration of proteins and ultra-fast MAS (> 40 kHz) lead to spectra with high proton resolution. The resulting proton chemical shift assignments as well as identification of long-range contacts between weakly coupled spins open up further possibilities in structure determination [11, 12].

Nuclear pore complexes

Nuclear pore complexes (NPCs) interconnect the cytoplasmic and intranuclear side of eukaryotic cells and control all nucleocytoplasmic exchange. Passage of molecules larger than 30 kDa usually is hindered unless they are able to bind to nuclear transport receptors (NTRs) [13]. Among the approximately 30 nucleoporins (Nups) that construct NPCs, the FG Nups at the inner surface of the central scaffold are known to be essential for building the permeability barrier. These proteins are enriched in phenylalanine-glycine dipetides and due to their high hydrophobicity are often considered to be unstructured. Although there has been made remarkable progress in understanding the architecture of NPCs and the identification of their molecular components [14], the NPC-mediated transport mechanism itself still remains elusive. Several models have been proposed to explain the main properties.

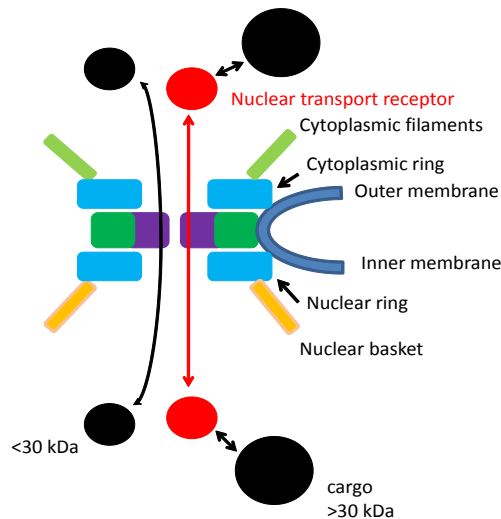


Figure 3: Sketch of the Nuclear Pore Complex (NPC). Rectangles represent different nucleoporins (Nups) filling up the inner part of the NPC. Same colors indicate that certain Nups are present on both sides of the cell. Small molecules can freely pass as well as larger material unless it is bound to Nuclear Transport Receptors (NTRs).

The virtual gate model assumes a polymer brush-like behavior of the FG Nups. Entropic motion of the elongated FG Nups excludes inert molecules that exceed a certain size whereas larger NTR cargo complexes can pass via binding to the FG repeat units [15]. The selective phase model is based on additional interactions between the FG domains (cohesiveness). They form a network whose mesh-size determines a limit for inert material. NTR cargo complexes can ‘melt’ through the network by dissociating successively nearby contacts [16]. FG Nup hydrogels exhibit properties similar to NPCs and provide a model to study and understand the functioning of NPCs [17]. Furthermore, elaborated studies speak in favor of the selective phase model [18].

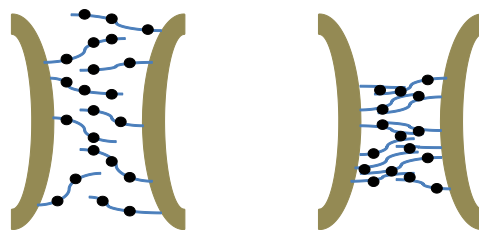


Figure 4: Schematic representation of the nature of the FG nucleoporins (Nups) within the NPC. Black circles denote FG motifs. According to the virtual gate model (left) the FG Nups exhibit polymer brush-like behavior. The selective phase model (right) assumes that they form a network via interactions between their FG domains (cohesiveness).

Polymers

Solid-state NMR applications have been used to study features of polymers and polymeric mixtures already for a long time in material science [19]. These studies include the investigation of general phenomena like exchange or confinement [20] as well as polymer specific behavior [21]. Recently, polymers also have gained importance in life-science applications. At first, a variety of polymer-based hydrogels has been used as biomaterials e.g. in contact lenses or as scaffold material in implants [22]. Nowadays, more advanced applications comprise the use of polymers and polymer assemblies that change their properties according to externally applied stimuli like pH or temperature [23]. Self-assembly of biocompatible polymers under defined conditions, that leads to micelle or hydrogel formation, has a huge potential as drug delivery systems. As a method which can probe different motional regimes, ssNMR is an excellent tool to study involved interactions by following the self-assembly process.

Sensory rhodopsin II

Sensory rhodopsin II (NpSRII) from *Natronomonas pharaonis* belongs to the family of microbial rhodopsins and is involved in signal transfer cascade across the membrane. It is a seven-helix (A-G) photoreceptor with retinal as a cofactor. The retinal is bound to a lysine residue (helix G) via a protonated Schiff base. Upon excitation with light the retinal undergoes an isomerization from all-*trans* to 13-*cis* configuration and induces conformational changes that activate the cognate transducer (NpHtrII). The NpHtrII is bound to the NpSRII by its transmembrane helices (TM1 and TM2). Its remaining part, including both HAMP domains (present in histidine kinases, adenylyl cyclases, methylaccepting chemotaxis proteins and phosphatases), is directed to the cytoplasmic side [24, 25]. The HAMP domains are believed to exist in a dynamic and compact conformation. The equilibrium between both states can be influenced by different factors (e.g. salt concentration and temperature). In lipid membranes, NpSRII/NpHtrII forms a 2:2 complex (see Figure 5). First structural information on NpSRII was obtained via X-ray crystallography [26]. Further structural investigation on NpSRII in micelles by solution [27] and membrane-embedded NpSRII by solid-state NMR followed [28]. In addition, ssNMR applications revealed dynamical features and water exposed residues of NpSRII as well as information regarding the interface of the NpSRII/NpHtrII complex after formation and activation [29].

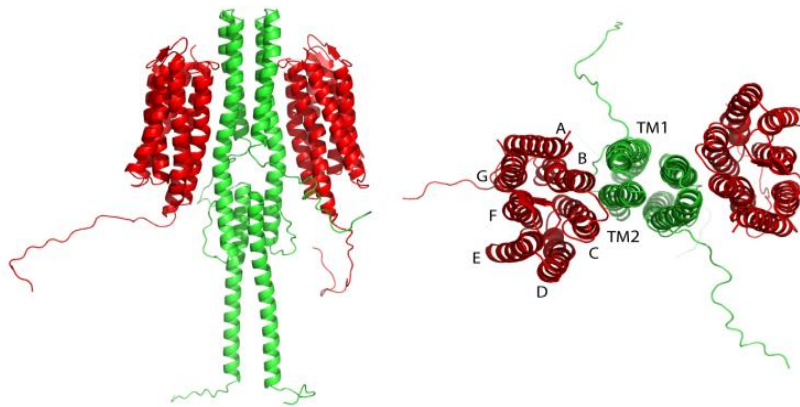


Figure 5: Sideview (left) and topview (right) of the NpSRII (red)/NpHtrII (green) complex. The presented structure is a homology model based on X-ray and ssNMR data [29].

Scope of the thesis

Solid-state NMR has become a powerful tool to investigate the structural organization and dynamical features of molecular assemblies such as micelles or hydrogels as well as a large variety of biomolecular samples. It allows to study insoluble and non-crystalline proteins which cannot be measured easily via other biophysical techniques as well as proteins in their native environment (e.g. membrane-embedded or in cellular settings). Specific labeling schemes and tailored NMR experiments have been developed to deal with biomolecular systems of increasing complexity. In addition, software programs that facilitate and automate the assignment process and application of new techniques like DNP [30, 31] open up further possibilities. In this context, the present thesis discusses further development of a software environment, ssNMR applications on different proteins and the impact of new approaches.

In **Chapter 2** we introduce the software environment and web portal FANDAS (Fast Analysis of multidimensional Nmr DATA Sets) that generates chemical shift predictions for standard NMR experiments in multiple spectral dimensions and facilitates the data analysis. Furthermore, the program can be used to make comparisons with reference data from other methods (e.g. X-ray crystallography) as well as to evaluate the efficiency of labeling schemes in an early stage of a planned (solid-state) NMR investigation.

Chapter 3 demonstrates in which way ssNMR methods can complete the overall picture with respect to other biophysical techniques in terms of structure and monitoring different motional regimes. The Nup98 FG nucleoporin of *Xenopus* NPCs and its O-glycosylated version are investigated in terms of hydrogel formation, transport properties, structure and dynamical behavior. In the end, the results are compared to former findings on the *S. cerevisiae* (sc) Nsp1p nucleoporin.

Chapter 4 discusses the structural and dynamical characterization of a protein hydrogel which emulates the sequence of the WT *S. cerevisiae* (sc) Nsp1p FG nucleoporin. The assignments and distance restraints which are obtained from dipolar-based ssNMR measurements are used for a structure calculation. The results in combination with information from scalar-based experiments allow to generate a first structural model.

Chapter 5 deals with a potential drug delivery system. Micelle formation of the thermosensitive amphiphilic copolymer mPEG-*b*-p(HPMAM-Nt₁₈-*co*-HPMAM-Lac₈₂) and the involved interactions is monitored by ssNMR experiments. In addition, drug loading capacities of the micelles are tested for the chemotherapeutic drugs paclitaxel and docetaxel.

Chapter 6 focuses on novel preparation and ssNMR methods to gain a further insight into the structure of the membrane protein Sensory Rhodopsin II and its receptor/transducer complex. Different labeling schemes and environments of the protein provide access to various structural aspects and information on the lipid-protein and protein-protein interface.

References

1. Renault, M., A. Cukkemane, and M. Baldus, *Solid-state NMR spectroscopy on complex biomolecules*. Angewandte Chemie International Edition, 2010. **49**(45): p. 8346-8357.
2. Renault, M., et al., *Cellular solid-state nuclear magnetic resonance spectroscopy*. PNAS, 2012. **109**(13): p. 4863-4868.
3. Andrew, E.R., A. Bradbury, and R.G. Eades, *Nuclear magnetic resonance spectra from a crystal rotated at high speed*. Nature, 1958. **182**(4650): p. 1659.
4. Pines, A., M.G. Gibby, and J.S. Waugh, *Proton-enhanced NMR of dilute spins in solids*. Journal of Chemical Physics, 1973. **59**(2): p. 569-591.
5. Andronesi, O.C., et al., *Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy*. Journal of American Chemical Society, 2005. **127**: p. 12965-12974.
6. Baldus, M., *ICMRBS founder's medal 2006: biological solid-state NMR, methods and applications*. Journal of Biomolecular NMR, 2007. **39**(1): p. 73-86.
7. Bloembergen, N., *Fine structure of proton magnetic resonance in $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$* . Physical Review, 1949. **75**: p. 1326.
8. Lange, A., S. Luca, and M. Baldus, *Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids*. Journal of American Chemical Society, 2002. **124**(33): p. 9704-9705.
9. Brunger, A.T., et al., *Crystallography and NMR System*. 1997, Yale University.
10. de Vries, S.J., M. van Dijk, and A.M. Bonvin, *The HADDOCK web server for data-driven biomolecular docking*. Nature Protocols, 2010. **5**(5): p. 883-897.
11. Ward, M.E., et al., *Proton-detected solid-state NMR reveals intramembrane polar networks in a seven-helical transmembrane protein proteorhodopsin*. Journal of the American Chemical Society, 2011. **133**(43): p. 17434-17443.
12. Reif, B., *Ultra-high resolution in MAS solid-state NMR of perdeuterated proteins: implications for structure and dynamics*. Journal of Magnetic Resonance, 2012. **216**: p. 1-12.
13. Peters, R., *Functionalization of a nanopore: The nuclear pore complex paradigm*. Biochimica et Biophysica Acta, 2009. **1793**: p. 1533-1539.
14. Alber, F., et al., *The molecular architecture of the nuclear pore complex*. Nature, 2007. **450**: p. 693-701.
15. Lim, R.Y.H., et al., *Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport*. PNAS, 2006. **103**(25): p. 9512-9517.
16. Ribbeck, K. and D. Görlich, *Kinetic analysis of translocation through nuclear pore complexes* EMBO 2001. **20**: p. 1320-1330.
17. Frey, S. and D. Görlich, *A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes*. Cell, 2007. **130**: p. 512-523.
18. Hülsmann, B.B., A.A. Labohka, and D. Görlich, *The Permeability of Reconstituted Nuclear Pores Provides Direct Evidence for the Selective Phase Model*. Cell, 2012. **150**: p. 738-751.
19. Schmidt-Rohr, K. and H.W. Spiess, *Multidimensional solid-state NMR and polymers*. 1994, San Diego: Academic Press Inc.
20. Saalwächter, K., *Glass-transition temperature gradient in nanocomposites: Evidence from nuclear magnetic resonance and differential scanning calorimetry*. Physical Review Letters, 2012. **108**(6): p. 065702(5).
21. Herrmann, A., et al., *Universal polymer dynamics revealed by field cycling ^1H NMR*. Macromolecules, 2009. **42**(14): p. 5236-5243.
22. Kopeček, J., *Hydrogels: From soft contact lenses and implants to self-assembled nanomaterials*. Journal of Polymer Science: Part A, 2009. **47**: p. 5929-5946.

23. de Graaf, A.J., et al., *A micelle-shedding thermosensitive hydrogel as sustained release formulation*. Journal of Controlled Release, 2012. **162**(3): p. 582-590.
24. Klare, J.P., et al., *The archaeal sensory rhodopsin II/transducer complex: a model for transmembrane signal transfer*. FEBS Letters, 2004. **564**: p. 219-224.
25. Klare, J.P., et al., *Effects of solubilization on the structure and function of the sensory rhodopsin II/transducer complex*. Journal of Molecular Biology, 2006. **356**(5): p. 1207-1221.
26. Luecke, H., et al., *Crystal structure of sensory rhodopsin II at 2.4 angstroms: insights into color tuning and transducer interaction*. Science, 2001. **293**(5534): p. 1499-1503.
27. Gautier, A., et al., *Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy*. Nature Structural & Molecular Biology, 2010. **17**(6): p. 768-775.
28. Etzkorn, M., et al., *Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy*. Angewandte Chemie International Edition, 2007. **46**(3): p. 459-462.
29. Etzkorn, M., et al., *Complex formation and light activation in membrane-embedded sensory rhodopsin II as seen by solid-state NMR spectroscopy*. Structure, 2010. **18**(3): p. 293-300.
30. Overhauser, A.W., *Polarization of nuclei in metals*. Physical Review, 1953. **92**: p. 411-415.
31. Rosay, M., et al., *Solid-state dynamic nuclear polarization at 263 GHz: spectrometer design and experimental results*. Physical Chemistry Chemical Physics, 2010. **12**(22): p. 5850-5860.

CHAPTER 2

Rapid prediction of multi-dimensional NMR data sets

Related publication: Sabine Gradmann, Christian Ader, Ines Heinrich, Deepak Nand, Marc Dittmann, Abhishek Cukkemane, Marc van Dijk, Alexandre M.J.J. Bonvin, Martin Engelhard and Marc Baldus, *J. Biomol. NMR*, **54**, 377-387 (2012)

Abstract

We present a computational environment for Fast Analysis of multidimensional NMR data sets (FANDAS) that allows assembling multidimensional data sets from a variety of input data and facilitates comparing and modifying such “*in-silico*” data sets during the various stages of the NMR data analysis. The input parameters can vary from (partial) NMR assignments directly obtained from experiments to values retrieved from *in-silico* prediction programs. The resulting predicted data sets enable a rapid evaluation of sample labeling in light of spectral resolution and structural content, using standard NMR software such as Sparky. In addition, direct comparison to experimental data sets can be used to validate NMR assignments, distinguish different molecular components, refine structural models or other parameters derived from NMR data sets. The method is demonstrated in the context of solid-state NMR (ssNMR) data obtained for the cyclic nucleotide binding domain of a bacterial cyclic nucleotide-gated channel and on membrane-embedded sensory rhodopsin II. FANDAS is freely available as web portal under WeNMR (<http://www.wenmr.eu/services/FANDAS>).

Introduction

NMR spectroscopy has become increasingly powerful to provide insight into the structural or dynamical organization of large biomolecules that cannot be obtained using other structural biology methods (see, e.g., [1-7]). Complexity in the resulting NMR spectra can not only result from the sheer molecular size giving rise to spectral crowding but it can also be related to the details of the experimental setup. For example, spinning sidebands are frequently observed in ^1H or ^{13}C edited solid-state NMR (ssNMR) experiments. Furthermore, the occurrence of polymorphism (see, e.g., [8, 9]) or the presence of several molecular components, such as in the case of cellular solid-state NMR studies [10] can complicate the spectral analysis.

To deal with such increasing levels of spectral complexity, tailored NMR approaches have been developed in the last years that facilitate different aspects of NMR-based analysis either in solution or in the solid state (see, e.g., [11, 12] for recent reviews). These methods often employ a combination of advanced sample preparation routines including specific labeling and spectroscopic filtering methods to simplify the resulting NMR data sets. In addition, significant progress has been made to support and automate the NMR data analysis process. Initially such approaches were developed for solution-state NMR applications (see, e.g., [13-17]). More recently, the underlying concepts have been

tailored for ssNMR studies [18-22]. Dedicated algorithms have also been proposed to aid other aspects of NMR studies, for example in the context of sample preparation [23] or homology modeling [24].

Next to NMR approaches, there is also a clear trend in combining information from various experimental sources and computations in order to tackle increasingly challenging systems. Methods such as cryo-electron microscopy (cryo-EM) or Small angle X-ray scattering (SAXS) are generating a wealth of lower resolution structural information that can be used to complement more classical structural studies by NMR and X-ray crystallography (see, e.g., [1]). All those data are being integrated with computational approaches for the modelling of large and complex systems [25]. Examples of such integrative approaches are HADDOCK [26, 27], an information-driven docking program that can incorporate various sources of experimental data and bioinformatics prediction to model macromolecular assemblies, and the Integrative Modeling Platform (IMP) which was designed for the modeling of macromolecular assemblies [28].

With these aspects in mind, the ability to directly relate structural information gathered by other biophysical methods to NMR parameters opens up new avenues to speed up different steps of an NMR-based analysis of molecular structure and motion. In such studies, a rapid comparison between reference data and NMR results during data analysis is often highly desirable. Here, we introduce the computational environment FANDAS (Fast Analysis of multidimensional NMR DAta Sets) that allows assembling multidimensional data sets from a variety of input data and to compare and modify such “*in-silico*” data sets during various stages of the NMR data analysis. We first discuss the general modules of FANDAS. Subsequently, we illustrate the use of FANDAS for the analysis of ssNMR data sets obtained on a cyclic nucleotide binding domain (CNBD) of a bacterial cyclic nucleotide-gated channel [29] and on membrane-embedded variants of sensory rhodopsin II (NpSRII). While the focus here is on solid-state NMR data, the underlying methodology can also be adapted for solution-state NMR studies. FANDAS is freely available as a web portal under WeNMR (<http://www.wenmr.eu/services/FANDAS>).

Material and Methods

Preparation of ¹³C labelled lipids from N. pharaonis

Natronomonas pharaonis were grown in a minimal medium (MM9) supplemented with 3.4 M NaCl, pH 9.2 and ¹³C-Na-acetate at 40°C while illuminating with orange light. After harvesting, the cells were washed with a buffer containing 250 g/L NaCl, 20 g/L MgSO₄ x 7 H₂O, 3 g/L sodium citrate and 2 g/L KCl. For lysis, 3 mg DNase A per L culture were added and the cell pellet was dialysed against double distilled H₂O (ddH₂O) over night. The lysate was sedimented at 170,000 x g, washed 3 times with ddH₂O and lyophilised. Lipids were extracted by treatment with CHCl₃/MeOH (1:1) under an argon atmosphere over night. The solvent was removed under reduced pressure; the extract was dissolved in CHCl₃/MeOH (1:1) and applied on a silica gel column equilibrated with diethyl ether/petroleum ether (85:15). The column was washed with diethyl ether/petroleum ether (85:15) and with CHCl₃/EtOH (2:1) containing 1% (v/v) ddH₂O and 1% (v/v) formic acid. CHCl₃/EtOH (1:1) with 10% (v/v) ddH₂O and 10% (v/v) formic acid was used to elute polar lipids. Lipid containing fractions were collected and the solvent was evaporated. In order to remove protein impurities, CHCl₃/MeOH (1:1) was added and the solution was centrifuged at 5000 x g, 10 min. This procedure was repeated until no protein impurities could be detected in a SDS-PAGE analysis. Freeze-dried ¹³C labeled *N. pharaonis* lipids were resuspended in 50 mM NaPi, pH 8 and after sonication stored at -80°C. ¹³C enrichment 99 % as determined by ESI mass spectrometry.

Protein expression

Uniformly ¹³C, ¹⁵N labeled NpSRII was expressed in *E. coli* and purified as described before ([30], [31]). Expression of NpSRII using *E. coli* cells grown in a minimal medium supplemented with 1,3-¹³C-glycerol was performed using the protocol of [39]. In both cases the yield was 3 mg/L medium. The N-terminally truncated transducer HtrII 157 was expressed in *E. coli* and purified as described in [32]. The proteins were reconstituted into unlabeled or ¹³C labeled lipids from *N. pharaonis* according to [33]. For the preparation of the CNBD domain, the reader is referred to [29]. The VLFY-reverse labeled sample was prepared as described in [30].

Experimental setup

NMR experiments were performed using Bruker Avance III spectrometers equipped with 3.2 mm triple resonance (¹H, ¹³C, ¹⁵N) MAS probeheads at 700 MHz (Figures 3-5), 500 MHz (Figure 6) and 800 MHz (Figure 7b) ¹H resonance frequency. The ¹³C-¹³C spin diffusion spectrum with a 20 ms mixing time on NpSRII was acquired at 4 °C effective temperature

and an MAS rate of 10.92 kHz. The 3D NCOCX experiment conducted on CNBD employed SPECIFIC-CP conditions [34] for sequential NCO transfer using radio frequency (r.f.) fields of 25 kHz and 42.5 kHz on ^{15}N and ^{13}C channels, respectively. The phase-switched MIRROR [35] scheme was applied for homonuclear (^{13}C , ^{13}C) mixing for a duration of 50 ms. The 3D data were acquired at 4 °C and a MAS rate of 17 kHz. The NMR measurements on the NpSRll/NpHtrll complex (Figure 6) were performed at an effective temperature of 5 °C and an MAS rate of 7.5 kHz. The 2D NCOCX experiment was acquired with a (^{13}C , ^{13}C) mixing time of 40 ms using proton-driven spin diffusion.

Results and Discussion

FANDAS Modules & Program use

Data creation within FANDAS generally proceeds using the steps depicted in Figure 1. The input data are generated in Step 1 of the program, based on the input amino acid sequence and secondary structure information if available. This procedure involves the creation of 31-tuples for each residue i of a given protein sequence that encodes proton chemical shifts (entries 1-15), backbone ^{15}N resonances (entry 16) and carbon chemical shifts (position 17-30). The nomenclature generally follows BMRB standards with some simplifications regarding the aliphatic CH_2 and CH_3 side-chain moieties (see supporting Table 1 for a complete list of definitions). NMR data sets of the protein of interest are hence represented by an $N \times 31$ dimensional matrix (with N corresponding to the amino acid sequence) of the form:

$$A_{i,j} = (\text{aa}_{i,0}, \text{cs}[X]_{i,j}) \text{ with } i = 1, \dots, N \text{ and } j = 1, \dots, 30.$$

In this matrix, the amino acid sequence aa is denoted in standard single letter format. In addition, it is also possible to include secondary structure information using an additional vector with the dimension N in which for each residue the single letters a , b or c/u stand for α -helix, β -strand or random coil/undefined, respectively (Figure 1, Step 1). Backbone chemical shift values (^1H , ^{15}N and ^{13}C) are taken from standard tables such as given in [36] corresponding to entries $\text{cs}[X]_{i,1-2}$, $\text{cs}[X]_{i,16}$ and $\text{cs}[X]_{i,17-19}$ (Table 1 in SI) and for side chains from the BMRB data base (<http://www.bmrb.wisc.edu/>). As shown in Figure 1, this information provides the input data that can subsequently be used for the prediction of NMR spectra.

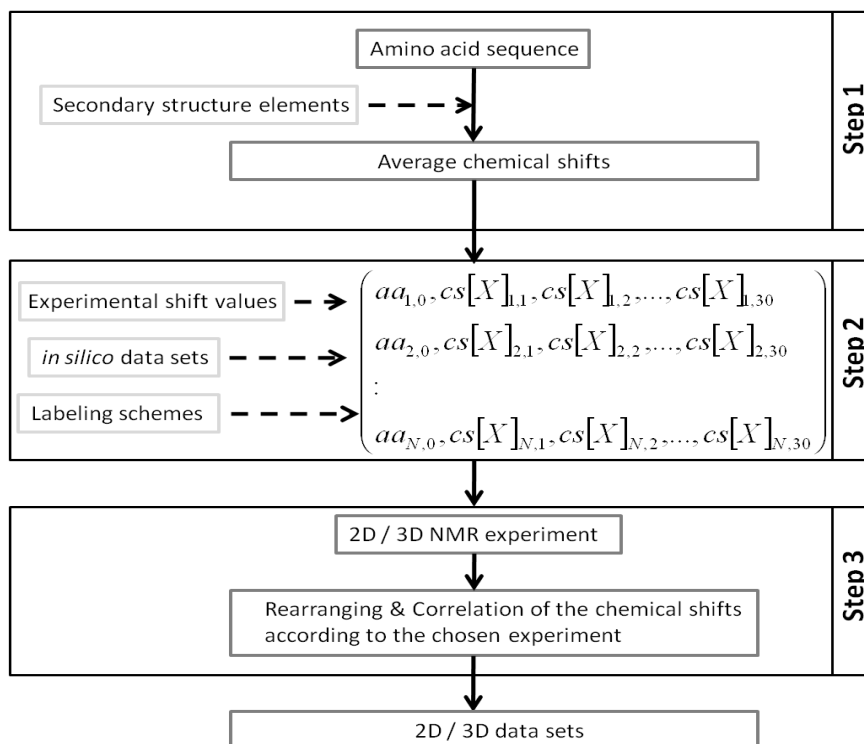


Figure 1: Flow chart of the different stages of FANDAS. Solid arrows denote compulsory steps and dashed lines reflect optional input modifications.

In Step 2 (see Figure 1), the input data can be further refined using experimental NMR chemical shift assignments deduced from reference samples or obtained under different experimental conditions. Furthermore, *in-silico* chemical shift data predicted by computational procedures such as ShiftX [37], ShiftX2 [38] or SPARTA+ [39] can be included. Furthermore, tailored labeling schemes can be used at the input stage 2 of FANDAS. So far, (1,3-¹³C)-glycerol labeling and (2-¹³C)-glycerol labeling [40-42], fractional deuteration [43, 44], and ¹³C, ¹⁵N reverse labeling [31, 45] have been implemented. Scrambling can be included as exemplified below. Finally, Step 2 can be used to provide distance constraints on the basis of a given PDB file and a user-defined distance cut off.

The resulting input matrix can be retrieved at every stage of the input session consisting of the stages 1 (*Input Generation*) and stage 2 (*Optional Input Modification*) (Figure 2). In the next stage (Step 3, *Spectrum prediction*), the user can select from a list of currently available 2D and 3D NMR experiments that can be predicted (Figure 2 and supporting Table 2). Figure 3 exemplifies the different stages of the input and spectrum modules for the case of the prediction of an intra-residue (¹³C, ¹³C) correlation spectrum assuming the peptide sequence Met-Ala-Thr-His-Glu (Figure 3a, upper panel) and 1,3-¹³C glycerol labeling (Figure 3a, lower segment). Since no information is given regarding the secondary

structure, the program assumes statistical average values for all atoms under consideration. Note that scrambling of individual atomic positions can be specified by selecting *indicate relative labeling* within the Spectrum prediction module (Figure 3b, upper segment). Furthermore, residue-specific probability limits can be defined allowing to adapt predictions to the expected amino-acid specific scrambling pattern. Depending on the experiment of interest, FANDAS rearranges and correlates the chemical shifts of the input file and generates predictions for spectra of intra-residue, sequential or distance-encoded data sets. For the later case, distance lists can be generated automatically from a PDB file under Step 2. In case of a multimer, FANDAS requires also the chain ID.

Finally, FANDAS returns a list of chemical shift predictions as an output file (Figure 3b, lower panel) that can be downloaded and subsequently used directly as a peak list in NMR analysis programs. For example, when using the Sparky software [15], the predictions are simply read in as a peak list via the command “rp”. Note that question marks in the output file have an important function as they indicate the dimension of the experiment. Atom names and further additional information (such as scrambling parameters or interatomic distances) are stored in the columns named “notes” of the Sparky peak list.

home >> FANDAS

FANDAS

WeNMR web portal

Return to the WeNMR website

WELCOME TO THE WENMR FANDAS WEB PORTAL >>

THE FANDAS WEB SERVER

Some small introduction to FANDAS

You have not yet created an input table from sequence. Proceed with step 1 or **retrieve** a previously saved file by name

STEP 1: INPUT GENERATION

Tool	Generated File
Generate input table from sequence & include average chemical shifts	

STEP 2: OPTIONAL INPUT MODIFICATION

Tool	Generated File
Read BMRB tables	
Read C-C and H-H distances from PDB	
Remove ¹³ C and ¹⁵ N shifts for reversely labeled amino acids	
Adapt input list to (1,3- ¹³ C)-glycerol labeling	
Adapt input list to (2- ¹³ C)-glycerol labeling	
Adapt input list to fractional deuteration	

STEP 3: SPECTRUM PREDICTION (SPARKY FORMAT)

Tool	Generated File
Predict intra-residue (¹³ C- ¹³ C) correlation spectrum	
Predict sequential (¹³ C- ¹³ C) correlation spectrum (i to i-1 and i+1)	
Predict (¹³ C- ¹³ C) correlation spectrum (distance encoded)	
Predict 2D DQ ¹³ C- ¹³ C spectrum (directly bonded C)	
Predict 2D ¹⁵ N- ¹³ CA spectrum	
Predict 2D ¹⁵ N- ¹³ CA- ¹³ CX spectrum (intra)	
Predict 2D ¹⁵ N- ¹³ CO spectrum	
Predict 2D ¹⁵ N- ¹³ CO- ¹³ CA/CB spectrum	
Predict 2D ¹⁵ N- ¹³ CO- ¹³ CX spectrum	
Predict backbone ¹ H- ¹⁵ N spectrum	
Predict ¹ H- ¹³ C spectrum (directly bonded)	
Predict ¹³ C- ¹ H spectrum (directly bonded)	
Predict (¹ H- ¹ H) NOESY spectrum (intra)	
Predict (¹ H- ¹ H) NOESY spectrum (distance encoded)	
Predict CHHC spectrum (distance encoded)	
Predict NHHC spectrum (distance encoded)	
Predict 2D CHH spectrum (distance encoded)	
Predict 2D NHH spectrum (distance encoded)	
Predict 2D HHC spectrum (distance encoded)	
Predict 3D HHC spectrum (distance encoded)	
Predict 3D CHH spectrum (distance encoded)	
Predict 3D NHH spectrum (distance encoded)	
Predict 3D ¹⁵ N- ¹³ CA- ¹³ CX spectrum (intra)	
Predict 3D ¹⁵ N- ¹³ CA- ¹³ CX spectrum (inter)	
Predict 3D ¹⁵ N- ¹³ CA- ¹³ CX spectrum (¹³ C distance encoded)	
Predict 3D ¹⁵ N- ¹³ CO- ¹³ CX spectrum	
Predict 3D ¹⁵ N- ¹³ CO- ¹³ CX spectrum (¹³ C distance encoded)	
Predict 3D ¹³ C S ² Q ² S ² Q spectrum (inter)	
Predict 3D ¹³ C D ² Q ² S ² Q spectrum (intra)	
Predict 3D ¹³ C D ² Q ² S ² Q spectrum (inter)	

Figure 2: Main page of the FANDAS program highlighting the different program steps as well as a list of currently available experiments under Step 3.

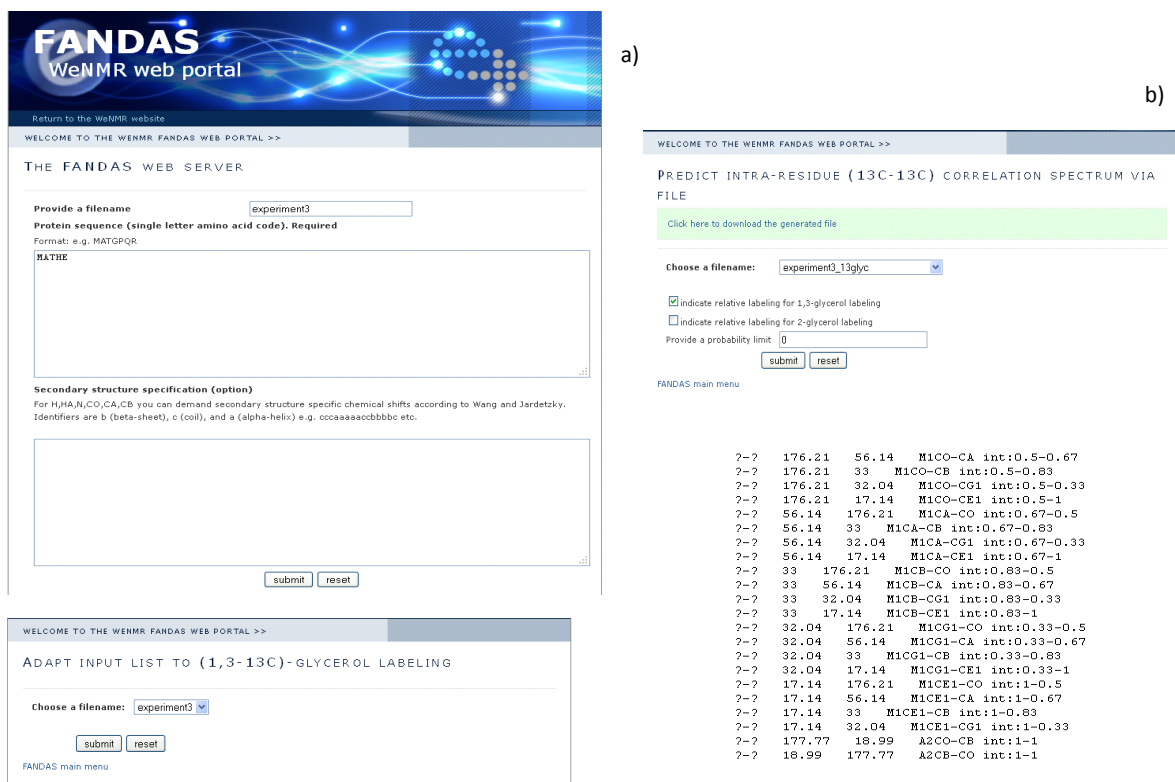


Figure 3: Interfaces reflecting the different stages of a FANDAS-based data production: In a, sequence input and definition of the file name (upper segment) as well as of the labeling pattern (1,3-¹³C glycerol labeling, lower panel) are shown. In b (upper panel), the experimental data set to be predicted (intra-residue CC, with a specific file name) as well as the relative labeling pattern with the certain probability limit is given. Finally, the lower panel of Figure 3b shows the resulting output file which can be read in as a peak list in NMR analysis software.

Applications

In the following, we demonstrate how FANDAS can facilitate different aspects of a ssNMR-based data analysis. In Figure 4a, a 2D ¹³C-¹³C proton-driven spin diffusion (PDSF) experiment is shown on sensory rhodopsin II (NpSRII) that was reconstituted using ¹³C labeled lipids from *N. pharaonis*. Data were acquired with a mixing time of 20 ms, leading to the dominant contribution of one-bond contacts in the spectrum. Since both protein and lipids are ¹³C labeled, the spectrum contains signal contributions of both molecular components. Protein chemical shift predictions by FANDAS help to distinguish lipid from the protein signals. Note that lipid signals are not only found in spectral regions well separated for aliphatic protein carbons, i.e., around 75-80 ppm but also extend into spectral segments between the 40 and 35 ppm region close to protein side chain correlations. Figure 4b shows an experimental ¹³C-¹³C spin diffusion spectrum of the pure

lipid sample with the same predictions confirming the FANDAS-based discrimination between protein and lipid signals on an experimental level.

In Figure 4, we made use of Step 2 of the FANDAS input module by combining *in-silico* and experimental data sets. Experimental NpSRII ssNMR chemical shift assignments (red) were taken from [31]. Additional chemical shift predictions were generated using ShiftX2 on the basis of the crystal structure (green, PDB file: 1GU8). Further modifications, for example by including solution-state NMR assignments are also possible. Average chemical shifts taken from the BMRB were included for side chain atoms using the simplified side chain topologies given in supporting Table 1 whenever experimental assignments were unavailable. Focusing on the threonine/serine region (Figure 4, insert) illustrates how FANDAS predictions can facilitate a further analysis of the experimental data.

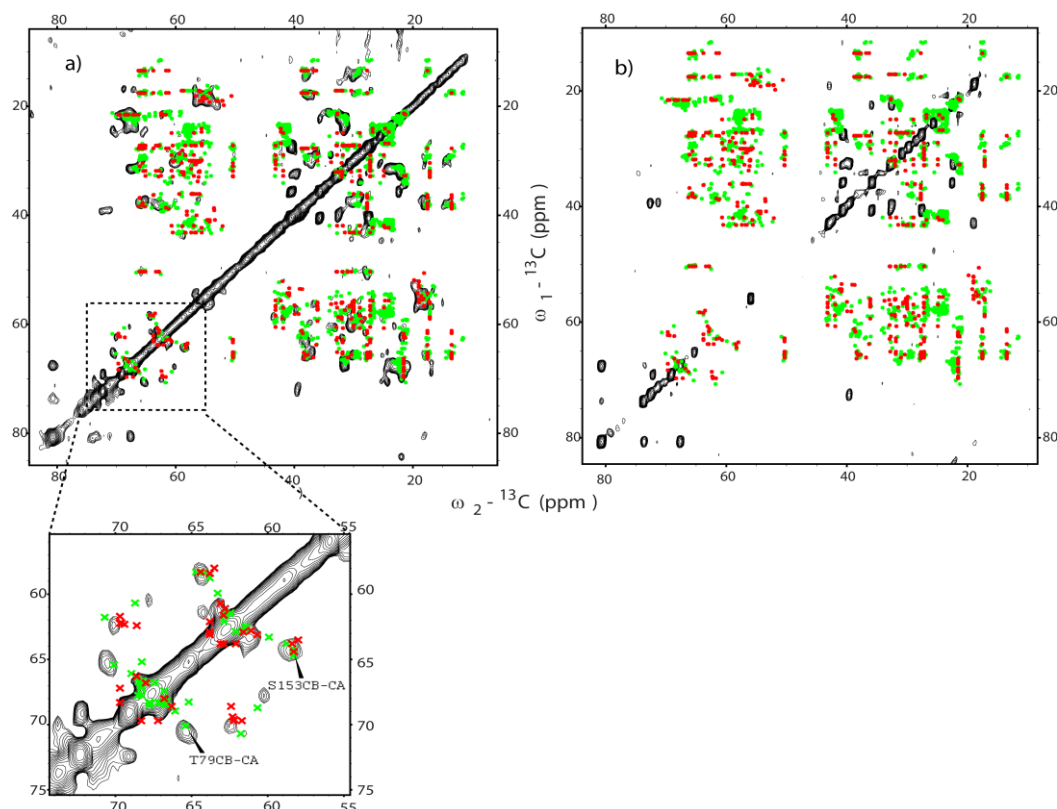


Figure 4: Comparison of experimental (^{13}C , ^{13}C) PDSO experiments on ^{13}C -labeled (NpSRII, lipids) (a) and ^{13}C liposomes (b). For the given mixing time of 20 ms, only intra-residue contacts are expected. FANDAS predictions (crosses) of the NpSRII were based on previous ssNMR assignments [31] (red) and the crystal structure (green; PDB file: 1GU8). Combination of those data facilitates the further analysis of the spectral region comprising threonine and serines (extract).

Here, C β -C α correlations corresponding to residue T79 as predicted from the X-ray structure (green) would be consistent with the assignment shown in the experimental

spectrum of Figure 4. Another example is S153 for which the FANDAS prediction (green) and the experimental assignment (red) nicely match.

In addition to the analysis of 2D data, FANDAS can also be used to predict three-dimensional data sets. In Figure 5, results are shown for an NCOCX experiment obtained on a (^{13}C , ^{15}N labeled) sample of the cyclic nucleotide binding domain (CNBD) [29] from a bacterial ion channel. FANDAS predictions were generated on the basis of solution-state NMR assignments [46] (Figure 5, red). In Figure 5a and b, consecutive planes in the 2-3 and 1-2 dimensions are shown, respectively. Comparison to the experimental data facilitated tentative assignments such as those indicated in green for a selected set of correlations. These were subsequently cross validated by additional ssNMR experiments leading to the assignments reported in [29]. Note that such studies can be performed interactively when scanning to the 3D experimental data sets.

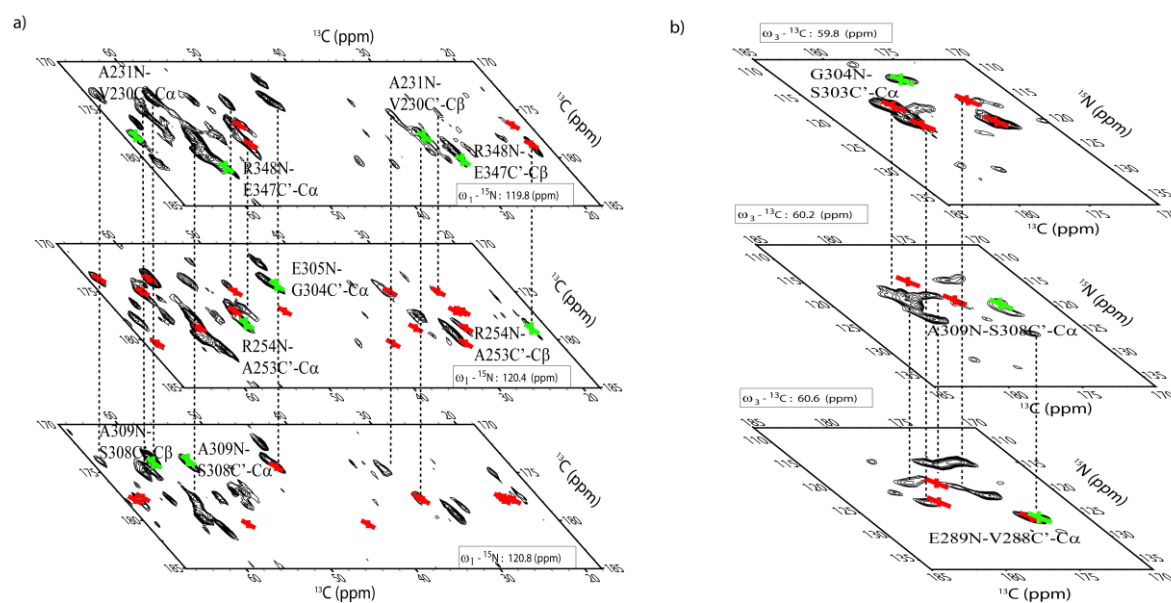


Figure 5: Consecutive planes of a 3D NCOCX experiment of CNBD along different axes. 2D planes indicated in the top and the lower panel are cut along the ^{15}N and $^{13}\text{C}\alpha$ frequencies, respectively. Chemical shift predictions (red) are based on solution state NMR assignments (PDB file: 2K06) and allowed to obtain several tentative assignments (green crosses).

Previously, sparse labeling has been used to obtain long-range distance constraints in several ssNMR studies [7, 42]. In such experiments, details of the protein labeling and potential scrambling effects can play an important role for the structural analysis. Again, FANDAS can aid such studies as demonstrated in Figure 6 for the case of the membrane-embedded NpSRII/NpHtrII complex. Here, only the receptor was isotope-labeled using 1,3- ^{13}C glycerol in the medium. The experimental results shown in Figure 6 correspond to a ^{13}C - ^{13}C spin diffusion spectrum with a mixing time of 150 ms. Chemical shift predictions

were generated for a distance-encoded ^{13}C - ^{13}C spectrum using 5 Å (red) and 3 Å (green) distance cut offs under the assumption of 1,3- ^{13}C glycerol labeling pattern. Chemical shift values as well as C-C distances were extracted from a structural model of NpSR_{II} in complex with NpHtr_{II} (1-157) obtained on the basis of crystallographic and NMR data [31, 47]. Comparison of the predictions to the experimental data indicates that several empirical correlations (such as cross peaks between 55 ppm and 15 ppm) must stem from polarization transfer over distances larger than 3 Å. The remarkable differences between the FANDAS spectra in Fig. 6a and b originate from the applied probability limit which is computed by multiplying the relative degree of ^{13}C labeling of spin X and spin Y, giving rise to cross peak XY with an intensity $L(X) \cdot L(Y)$ [48]. In Figure 6a, all non-vanishing correlations (i.e., $L(X) \cdot L(Y) > 0$) are included whereas Figure 6b assumes a limit $L(X) \cdot L(Y) > 0.55$. Such sets of predictions allow assessing the level of scrambling in a particular sample as long as structural information as well as a distance classification of the polarization transfer dynamics for a given mixing time are available [19, 49]. In the case of the NpSR_{II}/NpHtr_{II} sample, the FANDAS data could be used to refine the structural model and could be compared to additional experimental data obtained using longer (^{13}C , ^{13}C) mixing times.

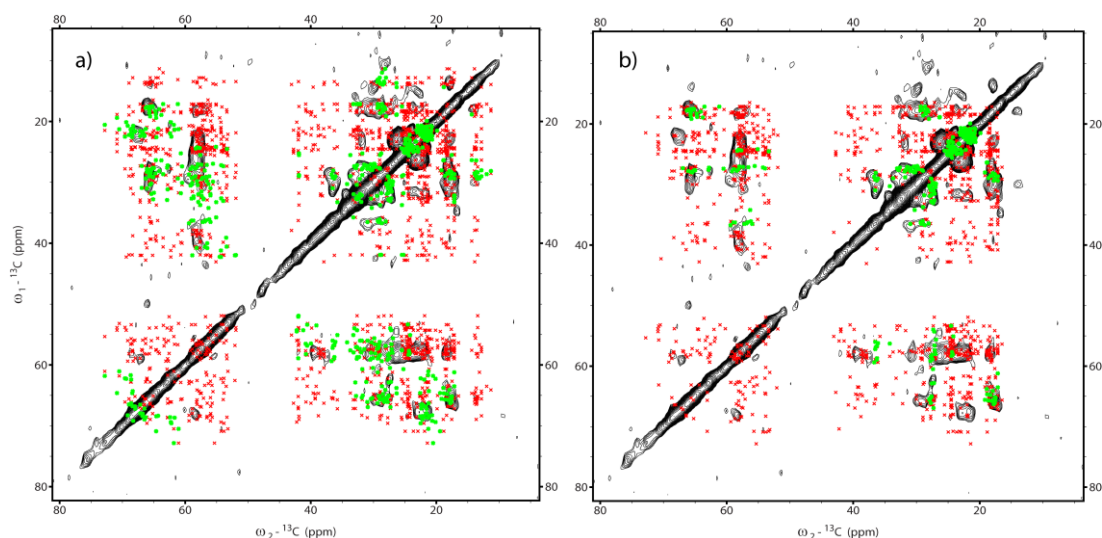


Figure 6: ^{13}C - ^{13}C spin diffusion (mixing time 150ms) spectrum of 1,3- ^{13}C glycerol labeled NpSR_{II} in complex with unlabeled NpHtr_{II} in lipid bilayers. Predictions were generated for a distance encoded ^{13}C - ^{13}C spectrum with 5 Å (red) and 3 Å (green) distance cut offs, respectively. Both, chemical shift values as well as distances were extracted from a homology model based on X-ray and NMR data [47]. In a), all possible correlations are present whereas in b) only correlations with a probability $\geq 56\%$ are included.

Finally, we demonstrate the utility of FANDAS in a stage where structural information available from other resources is used to decide on the optimal NMR labeling in terms of

spectral dispersion. As an example, we again consider the case of NpSRII. In Figure 7a, FANDAS predictions of an NCOCA spectrum were obtained using the crystal structure (PDB file: 1GU8) and assuming uniformly ^{13}C , ^{15}N labeled NpSRII (green) and a VLFY reverse labeled sample (red). FANDAS predicts a sizable reduction in spectral crowding, for example for ^{13}C resonance frequencies between 65 and 70 ppm. Taking into account the intrinsic accuracy of chemical shift predictions for ssNMR (see, e.g., [50]) and the possibility of N-C β correlations for an experimental PDSM mixing of 40 ms, experimental and predicted results (Figure 7b) are in good agreement. In a next stage, FANDAS could be used to aid an ssNMR-based analysis using complementary labeling schemes.

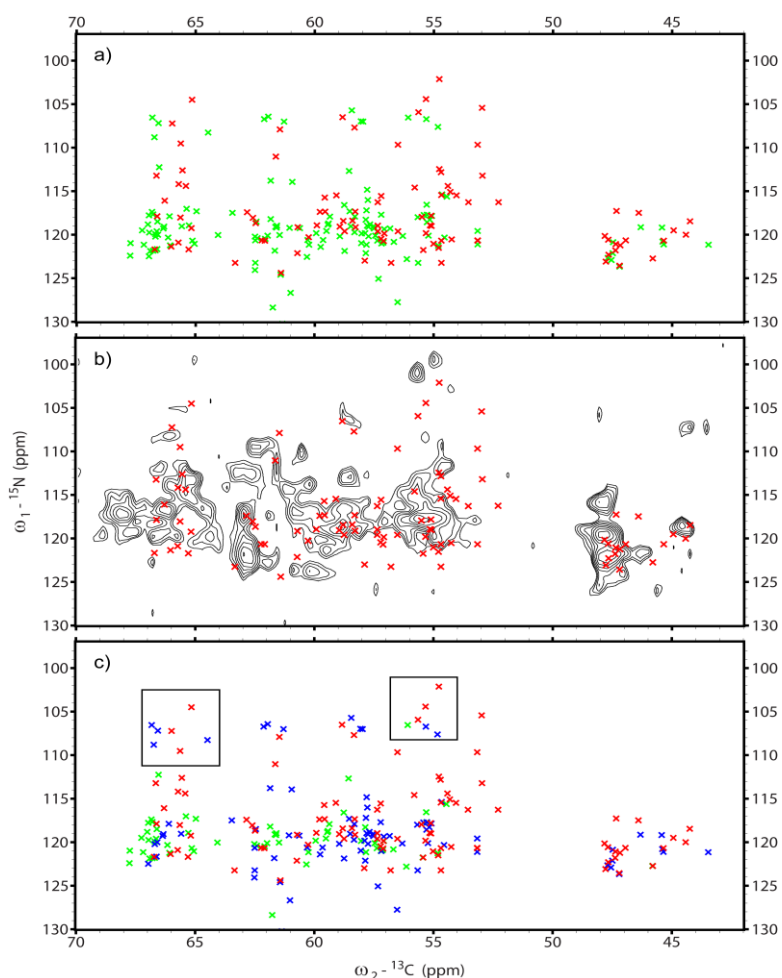


Figure 7: FANDAS predictions for a 2D NCOCA experiment for a ^{13}C , ^{15}N uniformly labeled NpSRII sample (green) as well as a VLFY (red) and TRIYM (blue, c) reverse labeled sample. Predictions are based on crystal structure (PDB file: 1GU8) suggesting a sizable reduction in spectral crowding and ambiguity for the reversed labeled preparations. In b, FANDAS predictions for a VLFY-labeled sample are compared to experimental results. In c, the two reverse labeling schemes are compared. Spectral regions in which FANDAS predicts complementary information, taking into account the intrinsic spectral resolution [50] are indicated by boxes.

As an example, Figure 7c compares predictions for fully labeled NpSRII (green), a VLFY reverse labeled sample (red) and a preparation of a TRIYM reverse labeled sample (blue). According to FANDAS, the combination of both labeling schemes would offer complementary spectral dispersion, for example in the indicated boxes referring to sequential contacts involving glycine-valine (blue) and glycine-isoleucine correlations (red) (^{13}C resonance frequency around 65 ppm) and glycine-leucine (blue) and glycine-alanine (red) close to the ^{13}C resonance frequency of 55 ppm.

Conclusion

We have introduced a computational environment that allows rapid prediction of multi-dimensional NMR data sets. These can be constructed even if empirical data are missing or incomplete and can be utilized at various stages of an NMR-based biomolecular study. Comparison of FANDAS sets to experimental results can facilitate the spectral assignment process as demonstrated for the case of 2D ^{13}C - ^{13}C spin diffusion experiments on NpSRII and a 3D NCOX experiment on CNBD. In addition, FANDAS readily allows distinguishing protein correlations from other molecular contributions such as lipids, peptidoglycans or nucleic acids [10, 51] and can be further extended to predict NMR correlations of such non-proteinaceous molecular components. Furthermore, labeling schemes and scrambling effects can be analyzed as illustrated here for the case of the membrane-embedded NpSRII/NpHtrII complex. Such considerations are important if distance restraints are to be deduced from multidimensional correlation experiments.

Especially in the early stage of an NMR study, the selection of the most suitable biomolecular labeling pattern or NMR experiment is often critical to maximize spectral resolution and dispersion. FANDAS can be readily used for such considerations because the availability of experimental data is not critical for its application. Since FANDAS works from chemical shift inputs derived from arbitrary PDB files, it also allows for an initial quality assessment of the NMR data if a reference structure is available. The modular character of FANDAS in principle allows further extensions in terms of (solid or liquid-state) NMR experiments, possibly including four-dimensional data sets or focusing on experiments that probe molecular motion on a residue-specific level. In summary, FANDAS provides a valuable tool to aid NMR-based studies of complex biomolecules at virtually all experimental stages starting from sample preparation to (supra)molecular structure elucidation.

References

1. Jehle, S., et al., *N-terminal domain of α B-crystallin provides a conformational switch for multimerization and structural heterogeneity*. Proceedings of the National Academy of Sciences, 2011. **108**(16): p. 6409-6414.
2. Lange, A., et al., *Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR*. Nature, 2006. **440**(7086): p. 959-962.
3. Ader, C., et al., *Amyloid-like interactions within nucleoporin FG hydrogels*. Proceedings of the National Academy of Sciences, 2010. **107**(14): p. 6281-6285.
4. Kato, H., et al., *Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR*. Proceedings of the National Academy of Sciences, 2011. **108**(30): p. 12283-12288.
5. Gelis, I., et al., *Structural Basis for Signal-Sequence Recognition by the Translocase Motor SecA as Determined by NMR*. Cell, 2007. **131**(4): p. 756-769.
6. Renault, M., et al., *Solid-State NMR on a Large Multidomain Integral Membrane Protein: The Outer Membrane Protein Assembly Factor BamA*. Journal of the American Chemical Society, 2011. **133**(12): p. 4175-4177.
7. Loquet, A., et al., *Atomic model of the type III secretion system needle*. Nature, 2012. **486**(7402): p. 276-279.
8. Seidel, K., et al., *High-resolution solid-state NMR studies on uniformly ^{13}C , ^{15}N -labeled ubiquitin*. ChemBiochem, 2005. **6**(9): p. 1638-1647.
9. Heise, H., et al., *Molecular-level secondary structure, polymorphism, and dynamics of full-length α -synuclein fibrils studied by solid-state NMR*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(44): p. 15871-15876.
10. Renault, M., et al., *Cellular solid-state nuclear magnetic resonance spectroscopy*. Proceedings of the National Academy of Sciences, 2012. **109**(13): p. 4863-4868.
11. Ruschak, A.M. and L.E. Kay, *Methyl groups as probes of supra-molecular structure, dynamics and function*. Journal of Biomolecular NMR, 2010. **46**(1): p. 75-87.
12. Renault, M., A. Cukkemane, and M. Baldus, *Solid-State NMR Spectroscopy on Complex Biomolecules*. Angewandte Chemie-International Edition, 2010. **49**(45): p. 8346-8357.
13. Nelson, S.J., D.M. Schneider, and A.J. Wand, *Implementation of the main chain directed assignment strategy - computer-assisted approach*. Biophysical Journal, 1991. **59**(5): p. 1113-1122.
14. Zimmerman, D.E., et al., *Automated analysis of protein NMR assignments using methods from artificial intelligence*. Journal of Molecular Biology, 1997. **269**(4): p. 592-610.
15. Goddard, T.D. and D. Kneller, *SPARKY 3*. 2006: University of California.
16. Linge, J.P., et al., *ARIA: automated NOE assignment and NMR structure calculation*. Bioinformatics, 2003. **19**(2): p. 315-316.
17. Guerry, P. and T. Herrmann, *Advances in automated NMR protein structure determination*. Quarterly Reviews of Biophysics, 2011. **44**(3): p. 257-309.
18. Fossi, M., et al., *SOLARIA: A protocol for automated cross-peak assignment and structure calculation for solid-state magic-angle spinning NMR spectroscopy*. Angewandte Chemie-International Edition, 2005. **44**(38): p. 6151-6154.
19. Manolikas, T., T. Herrmann, and B.H. Meier, *Protein Structure Determination from ^{13}C Spin-Diffusion Solid-State NMR Spectroscopy*. Journal of the American Chemical Society, 2008. **130**(12): p. 3959-3966.
20. Loquet, A., et al., *3D Structure Determination of the Crh Protein from Highly Ambiguous Solid-State NMR Restraints*. Journal of the American Chemical Society, 2008. **130**(11): p. 3579-3589.

21. Hu, K.N., W. Qiang, and R. Tycko, *A general Monte Carlo/simulated annealing algorithm for resonance assignment in NMR of uniformly labeled biopolymers*. Journal of Biomolecular NMR, 2011. **50**(3): p. 267-276.
22. Stevens, T., et al., *A software framework for analysing solid-state MAS NMR data*. Journal of Biomolecular NMR, 2011. **51**(4): p. 437-447.
23. Hefke, F., et al., *Optimization of amino acid type-specific ¹³C and ¹⁵N labeling for the backbone assignment of membrane proteins by solution- and solid-state NMR with the UPLABEL algorithm*. Journal of Biomolecular NMR, 2011. **49**(2): p. 75-84.
24. Brothers, M., et al., *VITAL NMR: using chemical shift derived secondary structure information for a limited set of amino acids to assess homology model accuracy*. Journal of Biomolecular NMR, 2011: p. 1-16.
25. Melquiond, A.S.J., et al., *Next challenges in protein-protein docking: From proteome to interactome and beyond.* WIREs Computational Molecular Science. WIREs Computational Molecular Science, 2012. **2**: p. 642-651.
26. Dominguez, C., R. Boelens, and A.M.J.J. Bonvin, *HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information*. J. Am. Chem. Soc., 2003. **125**(7): p. 1731-1737.
27. de Vries, S.J., M. van Dijk, and A.M.J.J. Bonvin, *The HADDOCK web server for data-driven biomolecular docking*. Nat. Protocols, 2010. **5**(5): p. 883-897.
28. Alber, F., et al., *Integrating diverse data for structure determination of macromolecular assemblies*, in *Annual Review of Biochemistry*. 2008, Annual Reviews: Palo Alto. p. 443-477.
29. Cukkemane, A., et al., *Solid-state NMR [¹³C,¹⁵N] resonance assignments of the Nucleotide-Binding Domain of a bacterial Cyclic Nucleotide-Gated Channel*. Biomol NMR Assign., 2012. **6**(2): p. 225-229.
30. Hohenfeld, I.P., A.A. Wegener, and M. Engelhard, *Purification of histidine tagged bacteriorhodopsin, pharaonis halorhodopsin and pharaonis sensory rhodopsin II functionally expressed in Escherichia coli*. FEBS Letters, 1999. **442**(2-3): p. 198-202.
31. Etzkorn, M., et al., *Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy*. Angewandte Chemie International Edition, 2007. **46**(3): p. 459-462.
32. Wegener, A.A., et al., *Structural insights into the early steps of receptor-transducer signal transfer in archaeal phototaxis*. EMBO Journal, 2001. **20**(19): p. 5312-5319.
33. Klare, J.P., et al., *Effects of solubilization on the structure and function of the sensory rhodopsin II/transducer complex*. Journal of Molecular Biology, 2006. **356**(5): p. 1207-1221.
34. Baldus, M., et al., *Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems*. Molecular Physics, 1998. **95**(6): p. 1197-1207.
35. Scholz, I., et al., *MIRROR recoupling and its application to spin diffusion under fast magic-angle spinning*. Chemical Physics Letters, 2008. **460**(1-3): p. 278-283.
36. Wang, Y. and O. Jardetzky, *Probability-based protein secondary structure identification using combined NMR chemical-shift data*. Protein Science, 2002. **11**(4): p. 852-861.
37. Neal, S., et al., *Rapid and accurate calculation of protein 1H, 13C and 15N chemical shifts*. Journal of Biomolecular NMR, 2003. **26**(3): p. 215-240.
38. Han, B., et al., *SHIFTX2: significantly improved protein chemical shift prediction*. Journal of Biomolecular NMR, 2011. **50**(1): p. 43-57.
39. Shen, Y. and A. Bax, *SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network*. Journal of Biomolecular NMR, 2010. **48**(1): p. 13-22.

40. LeMaster, D.M. and D.M. Kushlan, *Dynamical mapping of E-coli thioredoxin via ¹³C NMR relaxation analysis*. Journal of the American Chemical Society, 1996. **118**(39): p. 9255-9264.
41. Hong, M. and K. Jakes, *Selective and extensive ¹³C labeling of a membrane protein for solid-state NMR investigations*. Journal of Biomolecular NMR, 1999. **14**(1): p. 71-74.
42. Castellani, F., et al., *Structure of a protein determined by solid-state magic-angle- spinning NMR spectroscopy*. Nature, 2002. **420**(6911): p. 98-102.
43. Rosen, M.K., et al., *Selective methyl group protonation of perdeuterated proteins*. Journal of Molecular Biology, 1996. **263**(5): p. 627-636.
44. Nand, D., et al., *Fractional deuteration applied to biomolecular solid-state NMR spectroscopy*. Journal of Biomolecular NMR, 2012. **52**(2): p. 91-101.
45. Vuister, G.W., et al., *2D and 3D NMR-Study of Phenylalanine Residues in Proteins by Reverse Isotopic Labeling*. Journal of the American Chemical Society, 1994. **116**(20): p. 9206-9210.
46. Schunke, S., et al., *Solution structure of the Mesorhizobium loti K1 channel cyclic nucleotide-binding domain in complex with cAMP*. EMBO Rep, 2009. **10**(7): p. 729-735.
47. Etzkorn, M., et al., *Complex Formation and Light Activation in Membrane-Embedded Sensory Rhodopsin II as Seen by Solid-State NMR Spectroscopy*. Structure, 2010. **18**(3): p. 293-300.
48. Higman, V., et al., *Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively ¹³C-labelled proteins*. Journal of Biomolecular NMR, 2009. **44**(4): p. 245-260.
49. Lange, A., et al., *Analysis of proton-proton transfer dynamics in rotating solids and their use for 3D structure determination*. Journal of the American Chemical Society, 2003. **125**(41): p. 12640-12648.
50. Seidel, K., et al., *Comparative analysis of NMR chemical shift predictions for proteins in the solid phase*. Solid State NMR, 2009. **35**: p. 235-242.
51. Renault, M., et al., *Solid-State NMR Spectroscopy on Cellular Preparations Enhanced by Dynamic Nuclear Polarization*. Angewandte Chemie International Edition, 2012. **51**(12): p. 2998-3001.

CHAPTER 3

*Solid-state NMR characterization of functional hydrogels formed by *Xenopus Nup98* nucleoporins*

Related publication: Aksana A. Labohka, Sabine Gradmann, Steffen Frey, Bastian B. Hülsmann, Henning Urlaub, Marc Baldus and Dirk Görlich, *EMBO J.*, **32**, 204-218 (2013)

Abstract

Nuclear Pore Complexes (NPCs) consist of approximately 30 nucleoporins (Nups). Among them, the FG Nups play a crucial role for building the sieve-like permeability barrier in NPCs. Former studies showed that hydrogels comprising *S. cerevisiae* FG domains exhibited similar transport properties. Application of solid-state (ss) NMR measurements probed hydrophilic interactions between beta-sheets as essential for forming a functional gel. Here, the hydrogels of Nup98 and its O-glycosylated version from *Xenopus* were investigated. The O-GlcNAc modified hydrogel exhibited NPC-like properties whereas the non-glycosylated Nup98 was more restrictive as it blocked the entrance of large NTR cargo complexes. The higher mobility of the O-GlcNAc Nup98 was confirmed by dipolar-based ssNMR experiments that failed to obtain decent signal intensities after glycosylation. Furthermore, our studies revealed a complete absence of beta-sheet structures in the glycosylated but a high rigidity over the entire sequence and evidence for beta-strand formation in case of the non-glycosylated Nup98 FG hydrogel. Overall, this suggests that FG hydrogels from different organisms can assemble through several structural principles and acquire the same NPC-like permeability.

Introduction

Nuclear pore complexes (NPCs) connect the nuclear interior with the cytoplasm and control the exchange between the two compartments. They are built from nucleoporins (Nups; reviewed in [1, 2]) and are equipped with a sieve-like barrier that is freely permeable for small molecules, but becomes increasingly restrictive as inert mobile species approach or exceed 5 nm in diameter [3]. This limit corresponds to a mass of ≈ 30 kDa for spherical proteins. Nuclear transport receptors (NTRs) can overcome the size limit and transfer even very large cargoes across NPCs. Such “facilitated translocation” is used to supply nuclei with proteins and the cytoplasm with nuclear products such as ribosomes or mRNAs.

Facilitated translocation is remarkably efficient. A single NPC can accommodate ≈ 1000 transport events per second [4] and needs as little as ≈ 10 milliseconds to translocate an Imp β -cargo complex from the cytoplasmic to the nuclear side of the pore [5, 6]. These numbers imply that NPCs are able to translocate numerous objects simultaneously. Facilitated translocation is not directly coupled to an input of metabolic energy [7-9]. Instead it is primarily based on a higher permeability of NPC for NTRs as

compared to inert molecules; and it is the cooperation of importins or exportins with the RanGTPase system that renders NPCs into active and highly efficient cargo pumps [10].

Non-globular nucleoporin FG domains are central for the function of the NPC barrier [11-14]. They are typically several 100 residues long and comprise up to 50 FG, FxFG or GLFG motifs separated by spacers of rather variable sequence. For the sake of simplicity, these motifs are here collectively referred to as FG motifs. Metazoan NPCs contain several FG domains with multiple serine and threonine residues being modified with single O-linked β -N-acetylglycosamine (O-GlcNAc) moieties [15-18]. The function of this modification of FG domains is still unknown.

FG motifs bind NTRs during facilitated translocation [19-22]. How the FG-NTR interaction promotes NPC passage is, however, surprisingly difficult to answer. In fact, one would expect that a mere binding causes retention of NTRs and therefore slows down their passage. Also, such a simple binding cannot explain how inert material is excluded from passage. To resolve this paradox, several models have been proposed. The "virtual gate" model [23] assumes that brushes of FG domains repel inert material and that NTRs overcome this barrier by binding to the domains. The "selective phase" model [4] attributes an additional essential property to the barrier-forming FG domains, namely cohesiveness. It assumes that cohesive FG domains multivalently interact with each other and form a sieve-like hydrogel, the selective phase. The mesh size sets an upper size limit for unhindered NPC passage of inert material. According to the model, binding of an NTR to an FG motif disengages the corresponding repeat-repeat contact [4, 24]. NTR can thus "melt" through a dense FG hydrogel. The model is supported by the observations that the yeast *S. cerevisiae* (sc) contains several highly cohesive FG domains [13, 25-28], and that the FG domains from scNsp1p, the fused FG domains from scNup49p and scNup57p, or the human Nup153 FG domain indeed form FG hydrogels that display permeability properties very similar to authentic NPCs [25-27, 29]. That is, they exclude inert macromolecules > 5 nm, but allow an up to 20 000 times faster influx of the same macromolecule bound to an NTR. The rates of intra-gel diffusion are also consistent with the observed NPC transit times being in the 10 milliseconds range [5, 6].

Analysis of the scNsp1 FG domain revealed that hydrophobic as well as hydrophilic interactions contribute to inter repeat cohesion [27, 30]. Mutating the hydrophobic residues (which are mainly phenylalanines) to serines abolishes gel-formation. The N-terminal part of the domain (residues 1-175) is extremely cohesive and comprises inter

FG spacers that are very rich in Asn, Gln and Thr. Solid state NMR (ssNMR) revealed that these spacers engage in interchain β -sheets [30] - similar to those found in NQ-rich amyloids [31, 32]. The C-terminal part of the scNsp1 FG domain (residues \approx 274-601) is rather non-cohesive, apparently because the inter FG spacers are dominated by charged residues that counteract inter FG cohesion [28, 30]. It is, however, adhesive in the sense that it (weakly) binds to and greatly improves the selectivity of the NQ-rich scNsp1¹⁻¹⁷⁵ FG hydrogel.

If inter FG repeat cohesion were generally relevant for NPC function, then also eukaryotes other than *S. cerevisiae* should rely on this principle. We decided to test this prediction on the Nup98 FG hydrogel from the frog *Xenopus* and its O-GlcNAc-modified version. The NQ-content of Nup98 is low (12%) compared to scNsp1p (27%) but the highest of all *Xenopus* FG domains. Hydrogels were formed from both Nup98 versions and their transport properties were investigated via functional tests. Furthermore, application of ssNMR measurements revealed structural and dynamical features of both hydrogels in comparison to scNsp1p.

Experimental results

Hydrogel formation

Xenopus Nup98 has an N-terminal FG domain comprising 484 residues with an embedded GLEBS domain that interacts with the mRNA export mediator Gle2p/Rae1 [33] and it has the highest density of FG motives of all *Xenopus* FG domains, namely one motif per \approx 10 residues (see Figure 1b; GLEBS domain is underlined). We cloned and expressed the FG domain from Nup98 in *Escherichia coli*, purified it under conditions that prevent any premature cohesion, and concentrated it by lyophilization. We then dissolved the protein in an aqueous buffer at a concentration of \approx 200 mg/ml, which approximates the expected FG repeat concentration in authentic NPCs. Strikingly, the protein solution did not remain liquid but formed a gel.

We observed efficient hydrogel formation, however, also with a mutant lacking Asn and Gln altogether (Nup98 FG NQ \Rightarrow S; Figure 1a). It thus seems that *Xenopus* FG domains can form gels without the hydrophilic interaction potential of NQ-rich inter FG spacers. In this respect it is remarkable that *Xenopus* FG domains are considerably more hydrophobic than the FG domains from yeast scNsp1p or Nup100p. Thus, while inter FG cohesion in *S. cerevisiae* relies on both, hydrophobic and hydrophilic interactions [30], hydrogel formation in vertebrates appears to have a stronger hydrophobic basis.

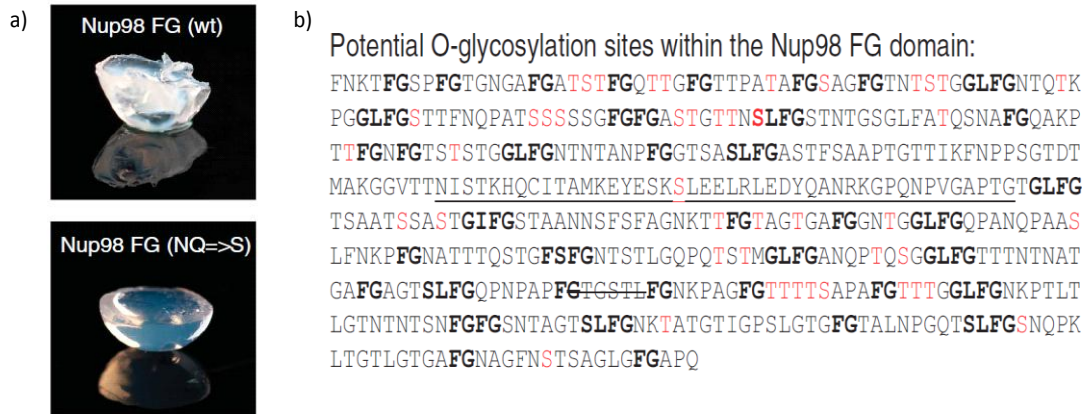


Figure 1: a) Macroscopic pictures of Nup98 (wild type; WT) and Nup98 (NQ => S mutant). b) Potential O-GlcNAc modification sites within the Nup98 FG domain were identified by mass spectrometry after enzymatic *in vitro* glycosylation. Modified Ser and Thr residues are shown in red, non-sequenced regions as strikethrough text, FG motifs in bold, the GLEBS domain is underlined.

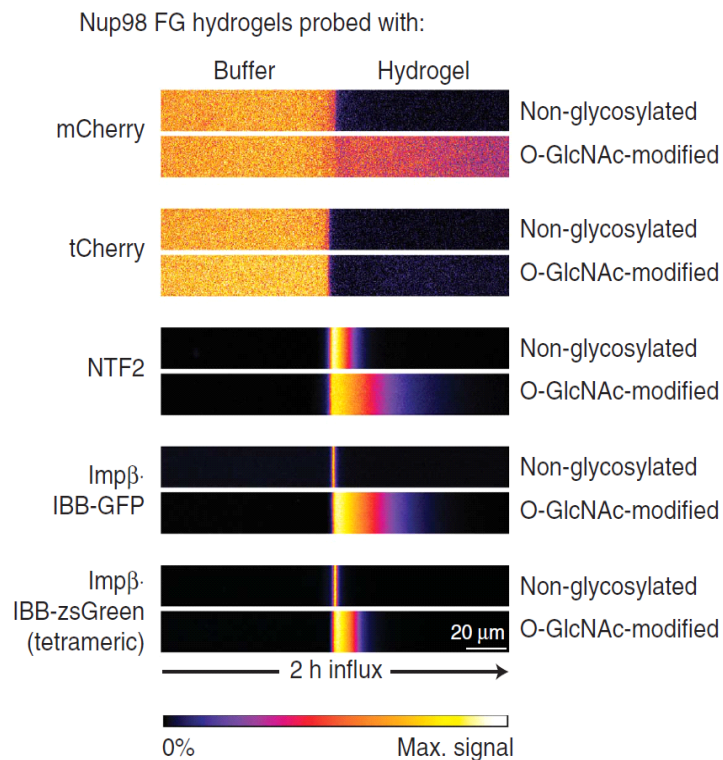


Figure 2: Non-glycosylated and O-GlcNAc-modified Nup98 hydrogels were probed with indicated mobile species.

Permeability properties

In a next step, we explored the permeability properties of Nup98 FG hydrogels. Specifically, we tested their ability to restrict the passage of inert molecules and to allow facilitated translocation of NTRs. For that, we formed the gels on microscope slides, equilibrated them in assay buffer, added fluorescent probes to the buffer reservoir, and measured the entry of these probes into the hydrogels by confocal laser scanning microscopy. The first set of probes included mCherry [34] as a ≈ 26 kDa inert molecule and Alexa488-labelled NTF2 as a small (29 kDa) NTR .

This probing revealed the Nup98 gel as a very selective barrier. The gel fully excluded mCherry, i.e. the gel/buffer partition coefficient for mCherry was essentially zero. The Nup98 gel is thus more restrictive than authentic NPCs, which allow GFP-sized molecules to pass at clearly detectable rates [3]. Nevertheless, NTF2 was able to enter the Nup98 gel efficiently, reaching within 30 min a partition coefficient of nearly 1000. Importantly, NTF2 did not remain stuck at the gel surface, but moved deep into the gel, on average 10 μm during 30 minutes incubation. At such diffusion rate, it would take ≈ 50 milliseconds to cross a 50 nm thick NPC barrier. This is ≈ 10 -fold slower than expected for passing an authentic NPCs [6], but consistent with the observation that the Nup98 FG hydrogel forms an extraordinary tight barrier. Closer inspection revealed, that the extreme tightness towards inert probes had an unexpected side effect, namely that all tested importins and exportins bound only to the gel surface and failed to visibly penetrate into the gel (shown for Imp β -cargo complexes in Figure 2). Such over-tight barrier in authentic NPCs would impede NTR-mediated nuclear transport of small proteins and even more so the transit of large (e.g. ribosome-sized) cargoes. We therefore considered several mechanisms that might confer a higher permeability to a Nup98 FG gel-based barrier.

A first possibility is that the local FG domain concentration of the *in vitro*-formed hydrogel is higher than in authentic NPCs, and so lowering the concentration should increase the permeability. The FG domain of e.g. scNsp1p forms a homogeneous gel over a wide range of concentrations (10-200 mg/ml). The Nup98 FG domain also forms a homogeneous gel at 200 mg/ml. When we attempted, however, to lower the concentration of the Nup98 FG domain, we observed a striking phase-separation into a protein-rich and a protein-poor phase. The microscopic appearance of the resulting gels then resembled a “holey Swiss cheese”. It should be noted that shrinking of the gel drop was evident also for a starting concentration of 200 mg/ml, though in this case, shrinking reduced the volume of

the entire gel drop and did not disintegrate the gel. The super-tight Nup98 gel therefore had a protein concentration probably exceeding 200 mg/ml.

The effect can be explained by the concept of a saturated FG hydrogel [25], which assumes that all cohesive units of an FG domain can find sufficiently close binding partners only when their concentration exceeds a certain saturation limit. Under-saturated gels should contain a significant share of un-paired cohesive units and hence have a larger mesh-size than the equivalent saturated gel. This state appears, however, unstable in an under-saturated Nup98 gel. Here it appears that the number of cohesive interactions is maximized by shrinking the gel volume until the saturation limit is reached. This shrinking then also leaves protein-poor areas behind (either by an increased buffer volume or as "holes in the cheese"). This effect qualifies the Nup98 FG domain as "highly cohesive". The effect precluded analyzing the permeability properties of a macroscopic, under-saturated Nup98 gel. It is well possible, however, that the anchoring of the domain to the rigid scaffold of an NPC would counteract such phase separation.

Another consideration is that native Nup98 is heavily modified by O-linked β -N-acetylglucosamines (O-GlcNAc; [35]), which are introduced by the O-GlcNAc transferase (OGT) that transfers GlcNAc moieties from UDP-GlcNAc to serines and threonines of numerous target proteins [36, 37]. To study the effects of this modification, we established a preparative method of enzymatically modifying the Nup98 FG domain. The reaction resulted in an electrophoretic size-shift of ≈ 9 kDa. Mass spectrometry revealed 46 potential modification sites, namely 29 threonines and 17 serines that occurred in clusters of up to 5 consecutively modified residues (Figure 1b). Otherwise, the modifications were distributed over the entire length of the FG domain. The *in vitro* reconstituted glycosylation reaction can be considered as a formal proof for Nup98 being an OGT substrate.

O-GlcNAc modification drastically changed the permeability properties of the Nup98 hydrogel. The glycosylated hydrogel allowed rapid entry and intra-gel diffusion not only of NTF2, but also of larger NTR-cargo complexes, such as Imp β -IBB-GFP (≈ 120 kDa) or a ≈ 480 kDa tetramerized Imp β -IBB-zsGreen complex (Figure 2). Concomitantly, the permeability toward mCherry increased. Nevertheless, the O-GlcNAc-modified Nup98 gel fully blocked entry of tCherry.

Structural characterization

As a next step, we analyzed uniformly ^{13}C , ^{15}N -labeled variants of the Nup98 gel by several solid-state NMR methods [38]. Through-space magnetization transfer (Cross-Polarization; CP) allows probing of very rigid regions that are stable for at least several milliseconds. Through-bond magnetization (Insensitive Nuclei Enhanced by Polarization Transfer; INEPT) detects highly mobile regions with motions in the nanoseconds regime. The two techniques are insensitive for protein segments moving in a time frame of micro- to milliseconds. However, this gap can be closed by direct excitation, which detects signals independently of mobility.

The previously characterized scNsp1 FG hydrogel yielded strong CP signals that can be attributed to the highly cohesive scNsp1¹⁻¹⁷⁵ FG subdomain (forming the rigid parts of the gel) as well as strong INEPT signals that predominantly stem from the non-cohesive and highly charged scNsp1²⁷⁴⁻⁶⁰¹ FG subdomain ([30]; Figure 3A and B). Analysis of the non-glycosylated Nup98 gel revealed very strong CP signals and very low INEPT signals. This is consistent with our observation that this type of gel is so rigid that importin β -type NTRs fail to enter. The almost complete lack of INEPT-visible mobile elements further suggests that the Nup98 FG domain is cohesive along its entire sequence. This assumption is supported by two-dimensional ^{13}C , ^{13}C experiments (see Figure 3C), indicating that all amino acid-specific cross-peaks detected by direct excitation also appear in the CP spectrum. It therefore appears that all NMR-resolved parts of the Nup98 FG domain reside in low mobility regions of the gel.

Based on these spectra, we conducted a qualitative analysis of secondary structure using secondary chemical shift information [30]. This analysis suggested that the rigid components of non-glycosylated Nup98 gel are not *per se* associated with the formation of β -strand structures as previously seen for the most rigid (NQ-rich amyloid-like) regions in scNsp1 FG hydrogel. This notion is based on the facts that the relevant cross-peaks in the Nup98 FG hydrogel are broader (Figure 3C, see below) and extend into spectral regions typically not associated with β -strands.

Upon glycosylation, CP signal intensities of the Nup98 gel dropped dramatically (Figure 3A), illustrating that the O-GlcNAc modification makes the gel less rigid. Interestingly, we observed only a tiny concomitant increase in the INEPT signal (Figure 3B), suggesting that the glycosylated gel was still largely devoid of fully mobile regions and that the O-GlcNAc-modified Nup98 FG domain remained cohesive along its entire sequence.

groups do not just prevent intragel contacts. Instead, they appear to engage in novel interactions that limit their mobility from a microseconds to milliseconds regime that seems characteristic for the entire glycosylated Nup98 gel.

Surprisingly, correlations compatible with β -sheet arrangements (Figure 3D, labeled in green) were completely absent from the O-GlcNAc-modified Nup98 gel. Instead, we observed correlations (Figure 3D, mustard yellow) that are typically observed in unstructured protein loop regions (also known as random coil correlations). This is in striking contrast to the interchain β -sheet structures that dominate inter FG repeat contacts of the highly cohesive NQ-rich FG domains from the yeast *S. cerevisiae* [30]. This points to a fully unanticipated conclusion, namely that FG hydrogels can assemble through different structural principles and yet acquire the same NPC-typical permeability.

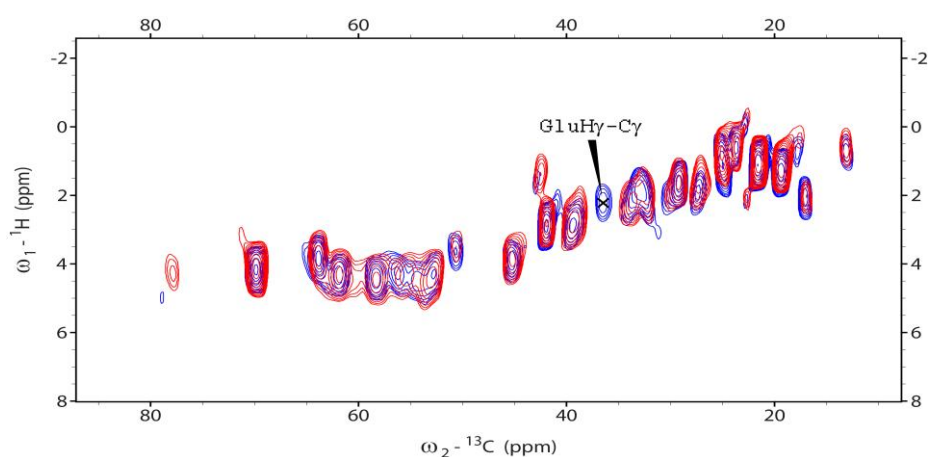


Figure 4: Overlay of scalar-based ^1H - ^{13}C spectra performed on the Nup98 FG hydrogel (blue) and the O-GlcNAc-modified version (red).

Despite the low INEPT signal, we were able to acquire a number of scalar-based ^1H , ^{13}C experiments of both hydrogels. These measurements allow to detect directly bond H-C correlations. In Figure 4 a spectral overlay of the non- (blue) and O-GlcNAc-modified (red) Nup 98 FG domain is shown. The extinction of the signal originating from the Glu Hy-Cy correlation is a further proof for a dynamical change after glycosylation. A glance at the Nup 98 sequence (Figure 1b) demonstrates that all Glu residues are located in the GLEBS domain, most of them in vicinity of a glycosylation site. Most likely, interactions emanating from the glycosylated site lead to a higher rigidity of this specific region.

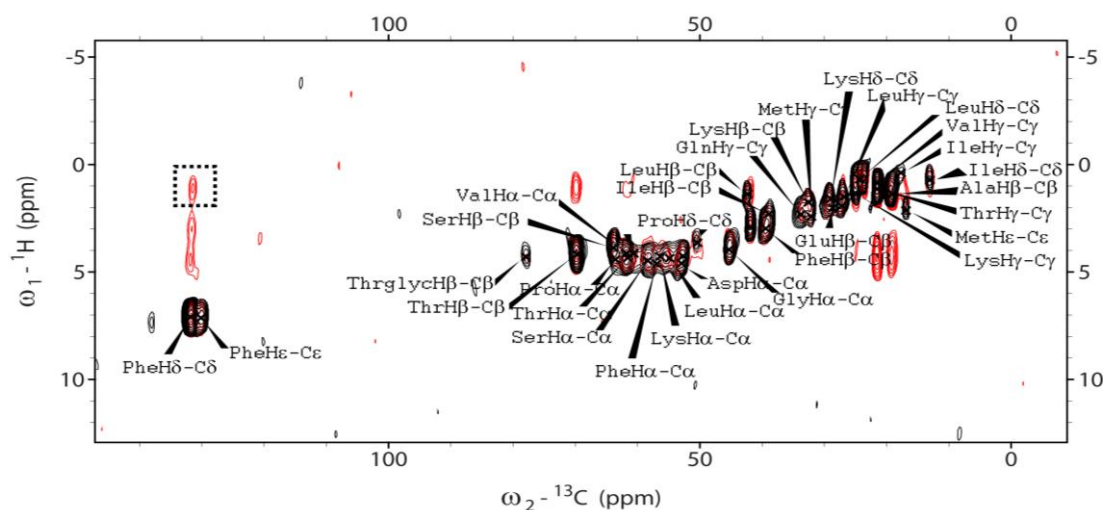


Figure 5: Overlay of a scalar-based ^1H - ^{13}C (black) and an $(^1\text{H}, ^1\text{H})$ - ^{13}C (red) experiment applied on the glycosylated Nup98 FG hydrogel. Signals originating from directly bond HC correlations are annotated. The dashed box indicates an inter-residue contact between Phe residues to the side chain region of Thr, Leu, Val and Ile.

In a further scalar-based ^1H , ^{13}C experiment an additional HH mixing step of 40 ms was included. During this mixing time magnetization is transferred via close-by protons. A comparison of an HHC (red) and HC (black) experiment of the O-GlcNAc-modified Nup98 gel is shown in Figure 5.

Signals originating from directly bond H-C correlations are annotated. Most of the additional peaks in the HHC spectrum can be attributed to intra-residue contacts. An exception is the signal highlighted by the dashed box. The correlation corresponds to a contact between Phe and the side chain region of other residues e.g. Thr or Leu. This might be an indication that the FG domains also in the O-GlcNAc-modified Nup98 to some extent are involved in hydrogel formation.

Conclusion and Discussion

The Nup98 FG domain as well as its O-GlcNAc-modified version formed hydrogels. Consequently, it is now evident that not only yeast contains cohesive FG domains [26, 28], but also *Xenopus*, i.e. an organism separated by ≈ 1 billion years of evolution. Furthermore, both hydrogels showed similar transport properties like real NPCs. However, the non-modified Nup98 carried out an extremely tight barrier which is more restrictive than in authentic NPCs. Glycosylation makes the entire gel more dynamic and

greatly favors the entry of large NTR-cargo complexes. The attached sugar is very hydrophilic and so it probably improves the water solubility of the hydrogel and allows the gels to retain a higher content of water. An interesting aspect is the reversibility of the O-GlcNAc modification [39]. Cells might exploit this effect and regulate the permeability of their NPCs through changes in the activities of the O-GlcNAc-transferase and/ or of the antagonistic glycosidase. Our preliminary data indicate that Nup98 from *Xenopus* eggs is O-GlcNAc modified to only 50% of the levels reached by excess of OGT, leaving indeed room for an upregulation as well as for downregulation of the modification. Structural investigation by ssNMR measurements revealed that β structures, which account for the most stable interactions at least within an scNsp1 FG hydrogel [28, 30, 40], are less prevalent in a non-modified Nup98 FG hydrogel and entirely absent in the fully O-GlcNAc-modified gel. It thus appears that the bulkiness of the sugars sterically excludes interchain β -sheets. The fact that the glycosylated Nup98 domain still forms a highly selective gel suggests that other intermolecular interactions dominate. One plausible structural element would be analogous to micelles, where hydrophobic residues of FG motifs aggregate to "miniaturized hydrophobic cores" that are interconnected by hydrophilic and possibly glycosylated spacers. The observation that the Nup98 FG domain has a considerably larger proportion of hydrophobic residues than the NQ-rich yeast FG domains is consistent with this assumption. Taken together, it appears that there are at least two fundamentally different ways of organizing a functional NPC barrier.

Why do vertebrates not rely on NQ-rich FG domains? One possible explanation is that such domains not only form reversible hydrogels, but also pose a risk of occasionally forming true amyloid fibers [40], which might then trigger pathogenic processes. The accumulated risk should be greater for long-lived vertebrates than for a unicellular, rapidly multiplying species.

Materials and Methods

Protein expression and purification

Xenopus FG and FG-like domains were predicted *in silico* from ESTs and the genome of *X. tropicalis*. The indicated protein sequences and domain boundaries refer to the following accession number (available from DDBJ/EMBL/GenBank): xtNup98: JX136847. Coding sequences were subcloned from IMAGE clones or cDNA into bacterial expression vectors. All proteins used were expressed in *Escherichia coli*, FG domains in the strain KY2266 [41],

and other proteins in BLR. The production of the following proteins has been described before: transportin, NTF2 [4], CRM1, RanQ69L¹⁻¹⁸⁰, PKI-NES-GFP [42], mCherry [26]. For all other proteins, new bacterial expression vectors were created.

The tetrameric Cherry was derived from mCherry by restoring the tetramerizing interface of DsRed [43].

All NTRs, transport substrates, and the OGT were purified by native Ni(II) chelate chromatography, Tev-cleavable tags were cleaved with Tev protease (with exception of His10-GFP-Tev-xIImp α). Proteins were further purified by gel filtration on a Superdex200 column equilibrated in 50 mM Tris/ HCl pH7.5, 200 mM NaCl, 2 mM DTT, and snap-frozen after adding 250 mM sucrose.

FG domains were essentially purified as described previously for the scNsp1 FG domain [25]. In brief, Ni-chelate chromatography was performed in 6 M Guanidinium hydrochloride. FG domains were then re-buffered to 20% acetonitrile + 0.08% TFA and lyophilized.

In vitro glycosylation

Ni(II) chelate beads were loaded with FG domains from bacterial lysates, washed with GuHCl buffer (50 mM Tris/HCl pH 7.5, 6 M Guanidinium hydrochloride, 2 mM DTT), and equilibrated in OGT-buffer (50 mM Tris/HCl pH7.5, 200 mM NaCl, 20 mM MgCl₂, 2 mM DTT, 1% Tween 20). The beads were then resuspended in 20 volumes of 5 μ M OGT and 1 mM UDP-GlcNAc (Sigma U4375) and rotated for 16h at room temperature. Beads were subsequently washed with OGT buffer, and GuHCl buffer and finally eluted with 0.5 M imidazole /HCl pH 7.5, 6 M guanidinium hydrochloride, 2 mM DTT. Eluted proteins were buffer-exchanged on a C18 column to 20% acetonitrile + 0.08% TFA and lyophilized.

FG hydrogel formation

Lyophilized FG domains were resuspended at 200 mg/ml in a buffer (see below). Gel formation was allowed for 16h at room temperature in a humidified chamber. Subsequently, the hydrogels were equilibrated in assay buffer (20 mM Tris pH7.5, 130 mM NaCl, 2 mM MgCl₂) for 16h at room temperature.

The FG domains studied here formed gels in a physiological buffer, however the handling was far easier when gelation was retarded by using 0.2 % TFA / 0.5-2 M Guanidinium hydrochloride as a resuspension buffer. All gels used in this study were formed under these conditions. To make the gels visible by fluorescence microscopy, they had been spiked with 0.2 μ M of Atto647N-labelled FG repeats.

Hydrogel permeation assays

Hydrogel permeation assays were performed as described [26] using a Leica SP5 confocal laser-scanning microscope equipped with a 63 x glycerol immersion objective.

NTRs and NTR-cargo complexes were used in assay buffer at a concentration of 1.5 μM (referring to monomeric substrates), while inert substances were used at 3 μM concentration.

Solid state NMR

Isotope-labeled, rich medium (110601402 *E.coli*-OD2 CN, Silantes, Mering, Germany) was used to produce the ^{13}C - and ^{15}N - double-labeled Nup98 FG domain. Purification and glycosylation of labeled protein was as described above. For gel formation, however, water was replaced by D_2O and gels were dialyzed against 200 mM Potassium phosphate buffer (pH 7.5) before analysis.

SsNMR experiments were conducted using 4 and 3.2 mm triple-resonance (^1H , ^{13}C , ^{15}N) probeheads at static magnetic fields of 11.7 T and 16.4 T (Bruker Biospin). Through-space transfer experiments were performed at 10.92 kHz magic angle spinning and 295 K. Typical proton field strengths for 90° pulses and SPINAL64 [44] decoupling ranged between 70 and 83 kHz. (^{13}C , ^{13}C) correlation spectra (e.g. Figure 7D) were obtained using proton-driven spin diffusion schemes under MIRROR conditions [45] employing mixing times of 40 ms. In the case of 2D ssNMR spectroscopy after direct excitation, mixing times were set to 20 ms (Nup98 FG) und 200 ms (O-GlcNAc-Nup98 FG). ^{13}C magnetization was produced using cross polarization (CP, contact times of 0.5 and 0.6 ms) or direct excitation using 90° ^{13}C pulses at 50 kHz r.f. field strength. Through-bond transfer experiments were performed at 10.92 kHz magic angle spinning and 295 K using an INEPT-TOBSY scheme [38]. A TOBSY [46] mixing time of 6.6 ms and 10 kHz GARP [47] decoupling were employed.

Mass spectrometric mapping of O-GlcNAc modified sites

The enzymatically glycosylated Nup98 FG domains was dissolved at 2 mg/ml in 2 M Urea, 50 mM Tris/ HCl pH7.5, 2 mM DTT and digested with 0.1 mg/ml chymotrypsin for 16h at 25°C . Peptides were desalted on a C18 matrix and subjected for to β -elimination (2h at 50°C and pH12), which converts a modified serine to an α -amino propenoic acid derivative and a modified threonine to an α -amino butenoic acid derivative [48]. The

resulting peptides were analyzed without further chemical modifications under standard conditions with an LTQ Orbitrap XL equipped with Agilent nano LC System. Obtained spectra were analyzed with Mascot 2.2 and searched for unmodified peptides and peptides containing the appropriate modifications. The non-modified FG domain was included as a negative control to sort out potentially false-positive glycosylation sites.

References

1. Brohawn, S.G., et al., *The nuclear pore complex has entered the atomic age*. Structure, 2009. **17**(9): p. 1156-1168.
2. Hetzer, M.W. and S.R. Wentz, *Border control at the nucleus: biogenesis and organization of the nuclear membrane and pore complexes*. Development Cell, 2009. **17**(5): p. 606-616.
3. Mohr, D., et al., *Characterisation of the passive permeability barrier of nuclear pore complexes*. EMBO, 2009. **28**(17): p. 2541-2553.
4. Ribbeck, K. and D. Görlich, *Kinetic analysis of translocation through nuclear pore complexes*. EMBO, 2001. **20**: p. 1320-1330.
5. Yang, W., J. Gelles, and S.M. Musser, *Imaging of single-molecule translocation through nuclear pore complexes*. PNAS, 2004. **101**(35): p. 12887-12892.
6. Kubitscheck, U., et al., *Nuclear transport of single molecules: dwell times at the nuclear pore complex*. Journal of Cell Biology, 2005. **168**(2): p. 233-243.
7. Schwoebel, E.D., et al., *Ran-dependent signal-mediated nuclear import does not require GTP hydrolysis by Ran*. Journal of Biological Chemistry, 1998. **273**(52): p. 35170-35175.
8. Englmeier, L., J.C. Olivo, and I.W. Mattaj, *Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis*. Current Biology, 1999. **9**(1): p. 30-41.
9. Ribbeck, K., et al., *The translocation of transportin-cargo complexes through nuclear pores is independent of both Ran and energy*. Current Biology, 1999. **9**(1): p. 47-50.
10. Görlich, D., M.J. Seewald, and K. Ribbeck, *Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation*. EMBO, 2003. **22**(5): p. 1088-1100.
11. Hurt, E.C., *A novel nucleoskeletal-like protein located at the nuclear periphery is required for the life cycle of Saccharomyces cerevisiae*. EMBO, 1988. **7**(13): p. 4323-4334.
12. Strawn, L.A., et al., *Minimal nuclear pore complexes define FG repeat domains essential for transport*. Nature Cell Biology, 2004. **6**(3): p. 197-206.
13. Patel, S.S., et al., *Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex*. Cell, 2007. **129**(1): p. 83-96.
14. Hülsmann, B.B., A.A. Labohka, and D. Görlich, *The Permeability of Reconstituted Nuclear Pores Provides Direct Evidence for the Selective Phase Model*. Cell, 2012. **150**: p. 738-751.
15. Finlay, D.R., et al., *Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores*. Journal of Cell Biology, 1987. **104**(2): p. 189-200.
16. Hanover, J.A., et al., *O-linked N-acetylglucosamine is attached to proteins of the nuclear pore. Evidence for cytoplasmic and nucleoplasmic glycoproteins*. Journal of Biological Chemistry, 1987. **262**(20): p. 9887-9894.
17. Holt, G.D., et al., *Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine*. Journal of Cell Biology, 1987. **104**(5): p. 1157-1164.
18. Onischenko, E.A., et al., *Cdk1 and okadaic acid-sensitive phosphatases control assembly of nuclear pore complexes in Drosophila embryos*. Molecular Biology of the Cell, 2005. **16**(11): p. 5152-5162.
19. Iovine, M.K., J.L. Watkins, and S.R. Wentz, *The GLFG repetitive region of the nucleoporin Nup116p interacts with Kap95p, an essential yeast nuclear import factor*. Journal of Cell Biology, 1995. **131**: p. 1699-1713.
20. Bayliss, R., et al., *Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP*. Journal of Molecular Biology, 1999. **293**(3): p. 579-593.
21. Rexach, M. and G. Blobel, *Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins*. Cell, 1995. **83**(5): p. 683-692.

22. Morrison, J., et al., *Solution NMR study of the interaction between NTF2 and nucleoporin FxFG repeats*. Journal of Molecular Biology, 2003. **333**(3): p. 587-603.
23. Rout, M.P., et al., *Virtual gating and nuclear transport: the hole picture*. Trends in Cell Biology, 2003. **13**(12): p. 622-628.
24. Kustanovich, T. and Y. Rabin, *Metastable network model of protein transport through nuclear pores*. Biophysical Journal, 2004. **86**(4): p. 2008-2016.
25. Frey, S. and D. Görlich, *A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes*. Cell, 2007. **130**(3): p. 512-523.
26. Frey, S. and D. Görlich, *FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties*. EMBO, 2009. **28**(17): p. 2554-2567.
27. Frey, S., R.P. Richter, and D. Görlich, *FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties*. Science, 2006. **314**(5800): p. 815-817.
28. Yamada, J., et al., *A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins*. Molecular & Cellular Proteomics, 2010. **9**(10): p. 2205-2224.
29. Milles, S. and E.A. Lemke, *Single molecule study of the intrinsically disordered FG-repeat nucleoporin 153*. Biophysical Journal, 2011. **101**(7): p. 1710-1719.
30. Ader, C., et al., *Amyloid-like interactions within nucleoporin FG hydrogels*. PNAS, 2010. **107**(14): p. 6281-6285.
31. Balbirnie, M., R. Grothe, and D.S. Eisenberg, *An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated beta-sheet structure for amyloid*. PNAS, 2001. **98**(5): p. 2375-2380.
32. Eanes, E.D. and G.G. Glenner, *X-ray diffraction studies on amyloid filaments*. Journal of Histochemistry and Cytochemistry, 1968. **16**(11): p. 673-677.
33. Pritchard, C.E., et al., *RAE1 is a shuttling mRNA export factor that binds to a GLEBS-like NUP98 motif at the nuclear pore complex through multiple domains*. Journal of Cell Biology, 1999. **145**(2): p. 237-254.
34. Shaner, N.C., et al., *Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein*. Nature Biotechnology, 2004. **22**(12): p. 1567-1572.
35. Powers, M.A., et al., *Reconstituted nuclei depleted of a vertebrate GLFG nuclear pore protein, p97, import but are defective in nuclear growth and replication*. Journal of Cell Biology, 1995. **128**(5): p. 721-736.
36. Haltiwanger, R.S., G.D. Holt, and G.W. Hart, *Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase*. Journal of Biological Chemistry, 1990. **265**(5): p. 2563-2568.
37. Lubas, W.A. and J.A. Hanover, *Functional expression of O-linked GlcNAc transferase. Domain structure and substrate specificity*. Journal of Biological Chemistry, 2000. **275**(15): p. 10983-10988.
38. Andronesi, O.C., et al., *Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy*. Journal of the American Chemical Society, 2005. **127**(37): p. 12965-12974.
39. Gao, Y., et al., *Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain*. Journal of Biological Chemistry, 2001. **276**(13): p. 9838-9845.
40. Halfmann, R., et al., *Prion formation by a yeast GLFG nucleoporin*. Prion, 2012. **6**(4): p. 391-399.

41. Kanemori, M., et al., *Synergistic roles of HslVU and other ATP-dependent proteases in controlling in vivo turnover of sigma32 and abnormal proteins in Escherichia coli*. Journal of Bacteriology, 1997. **179**(23): p. 7219-7225.
42. Güttler, T., et al., *NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1*. Nature Structural & Molecular Biology, 2010. **17**(11): p. 1367-1376.
43. Wall, M.A., M. Socolich, and R. Ranganathan, *The structural basis for red fluorescence in the tetrameric GFP homolog DsRed*. Nature Structural Biology, 2000. **7**(12): p. 1133-1138.
44. Fung, B.M., A.K. Khitritin, and K. Ermolaev, *An improved broadband decoupling sequence for liquid crystals and solids*. Journal of Magnetic Resonance, 2000. **142**(1): p. 97-101.
45. Scholz, I., et al., *MIRROR recoupling and its application to spin diffusion under fast magic-angle spinning*. Chemical Physics Letters, 2008. **460**: p. 278-283.
46. Baldus, M. and B.H. Meier, *Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity*. Journal of Magnetic Resonance A, 1996. **121**: p. 65-69.
47. Shaka, A.J., P.B. Barker, and R. Freeman, *Computer-optimized decoupling scheme for wideband applications and low-level operation*. Journal of Magnetic Resonance, 1985. **64**: p. 1567-1572.
48. Wells, L., et al., *Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications*. Molecular & Cellular Proteomics, 2002. **1**(10): p. 791-804.

CHAPTER 4

Structural investigation of a homopolymeric protein hydrogel via solid-state NMR

Sabine Gradmann, Steffen Frey, Dirk Görlich and Marc Baldus, Manuscript in preparation

found to contain beta-sheet structural elements [6].

Moreover, intermolecular beta-sheet contacts were essential for forming the hydrogel. The charged C-terminal of the protein rich in FSFG units is very flexible and does not have a defined structure. Both parts were substantial for establishing a permeability barrier similar to real NPCs [6]. Based on these findings a homopolymeric version of the WT scNsp1 hydrogel was constructed that showed similar transport properties in functional tests. In the following this protein construct will be referred to as HomoNsp1. It comprises in total 350 residues with a 20 amino acid long repeat unit (one strand is highlighted in italic and bold letters in Figure 1). This repeat unit contains one FG motif (FSFG) surrounded by spacers enriched in Asn, Thr and Ser. Consequently, it is a sequence with the main characteristics of the WT scNsp1 FG domain (NTS rich part and FSFG unit).

In the present study, we investigated the dynamical and structural features of this hydrogel by ssNMR measurements.

Results and Discussion

Gel Dynamics and Mobility as seen by ssNMR

To analyze the overall dynamical behavior of the protein we performed ssNMR experiments with dipolar-based (cross polarization, CP) and scalar-based (insensitive nuclei enhanced by polarization transfer, INEPT) magnetization transfer to probe the rigid and mobile parts of the sample, respectively [7]. Figure 2 compares the 1D ^{13}C spectra for the two different types of transfer (black: CP, red: INEPT). Apparently, the sensitivity is similar in both experiments. In contrast, distinct differences become apparent when comparing carbonyl (175 ppm) and aromatic regions (140-120 ppm) of the spectra. In the case of scalar-based transfer (Figure 2, red), carbonyl signals are largely absent and the aromatic region of the ^{13}C spectrum is dominated by sharp resonances that we assign partially to His correlations of the N-terminal tag. The changes in the aliphatic region of the spectrum, in particular around ^{13}C resonances of 30 ppm could also be attributed to the Lys (C δ) and His (C β) residues in the N-terminal tag. The signal around 40 ppm corresponds to Asn (C β).

For a further amino-acid specific analysis we performed two-dimensional ssNMR experiments using through-space (dipolar-based) and through-bond (scalar-based) transfer schemes. Remarkably, the corresponding 2D spectra (Figure 3) indicate that most residue types are visible in both spectra. Only the His and Lys residues of the His-Tag are

present exclusively in the INEPT-TOBSY [8] spectrum in line with our observations in 1D (Figure 2). However, the signal intensities of the INEPT-TOBSY spectrum cannot be mapped entirely on the amino acid distribution of the His-Tag e.g. correlations of Asn residues are visible in the spectrum.

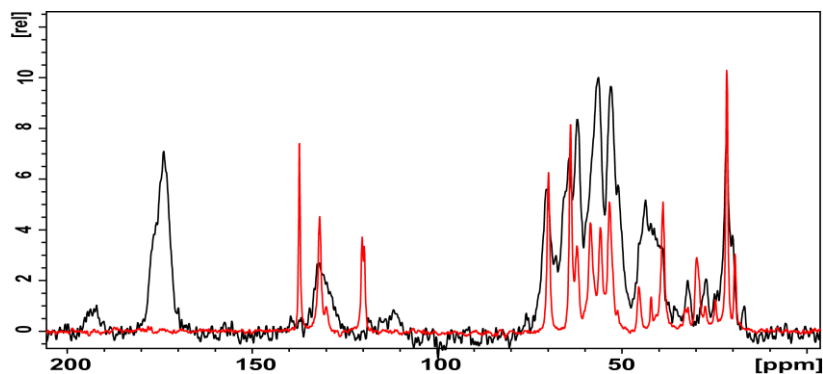


Figure 2: Comparison of 1D ^{13}C spectra probing rigid (black) and mobile parts (red) of the protein hydrogel.

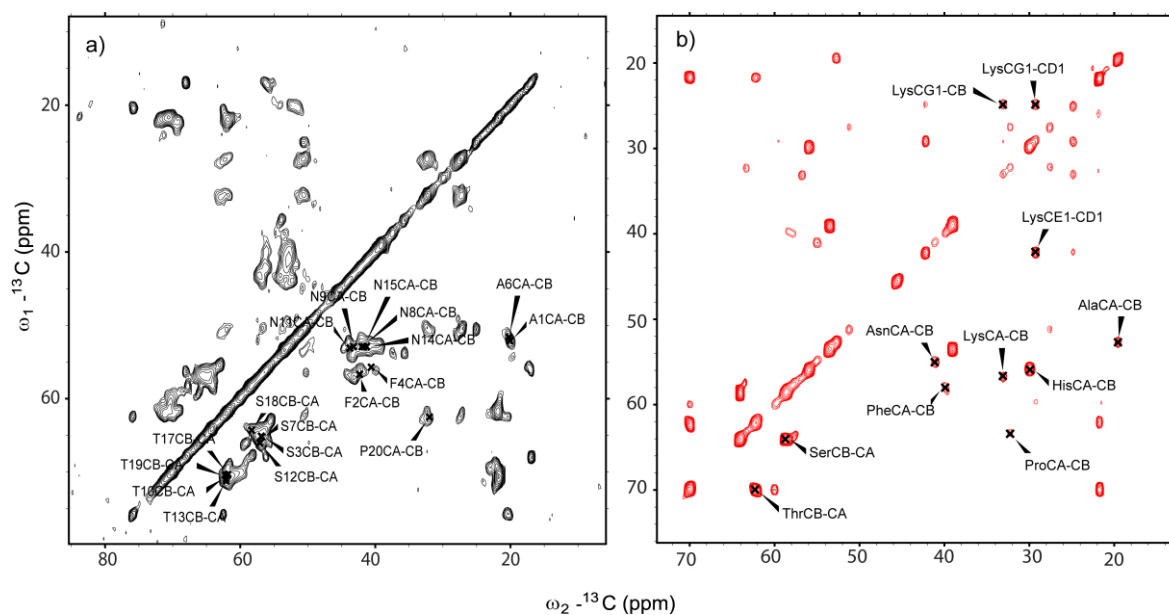


Figure 3: 2D (^{13}C , ^{13}C) analysis of a) rigid (black) and mobile parts (red) of the protein hydrogel HomoNsp1. a) corresponds to a ^{13}C - ^{13}C PDSF and b) an ^{13}C - ^{13}C INEPT-TOBSY experiment. Identified intra-residue contacts are annotated.

Consequently, the mobile contributions do not stem from a mobile His-Tag alone. A comparison of the relative intensities in the INEPT-TOBSY and PDSF spectrum indicates a similar intensity distribution over the residues of the sequence (see Table 1). In both cases, the residues with the highest occurrence per repeat unit (Asn, Ser, Thr) also show

the largest intensities. In contrast, ssNMR intensities for the remaining residues are considerably lower.

Amino acid	N-terminus	INEPT-TOBSY Intensity	Repeat unit	PDSD Intensity
N	0	9.9	5	4.8
S	5	12	4	10
T	2	9.5	4	6.2
A	1	0.93	2	2.3
F	0	0.7	2	3.1
P	0	0.55	1	2.7
H	14	3.8	0	0
K	1	0.96	0	0

Table 1: Intensity distributions of 2D INEPT-TOBSY and PDSD spectra.

This suggests that the hydrogel contains whole repeat units of high and low mobility. To estimate their relative influence across the entire protein sequence, we compared ssNMR intensities with amino acid occurrences in the sequence using the His tail as a reference. Back calculation of INEPT-TOBSY intensities gives a value of 3.8 for 14 His residues. An intensity of 9.9 for Asn signals would correspondingly suggest that around 36 Asn residues of the 80 in the whole sequence contribute to mobile ssNMR signals. Hence, we estimate an approximate distribution of 50 % mobile and rigid components within the hydrogel.

Resonance Assignments and Secondary Structure Analysis

So as to obtain (^{13}C , ^{15}N) chemical shift assignments we performed a number of 2D dipolar-based experiments. (^{13}C , ^{13}C) proton-driven spin diffusion (PDSD) with short (Figure 4b) and long mixing times (4a), NCOCX (4c) and NCACX (4d) experiments to obtain intra- and inter-residue information [9, 10] are given below.

In a first step, we tried to assign the CP based data on the basis of one repeat unit from the protein sequence (see highlighted part in Figure 1). As Figure 4 shows nearly all signals in the spectra can be explained by backbone resonances of one consecutive repeat unit. Consequently, all repeat units in the protein sequence which are detected by the dipolar-based experiments must exhibit very similar chemical shifts. The assignment strategy is exemplified on the unique residue combination T10-N11-S12. Inter-residue peaks corresponding to S12-N11 and T10-N11 correlations were identified in the PDSD

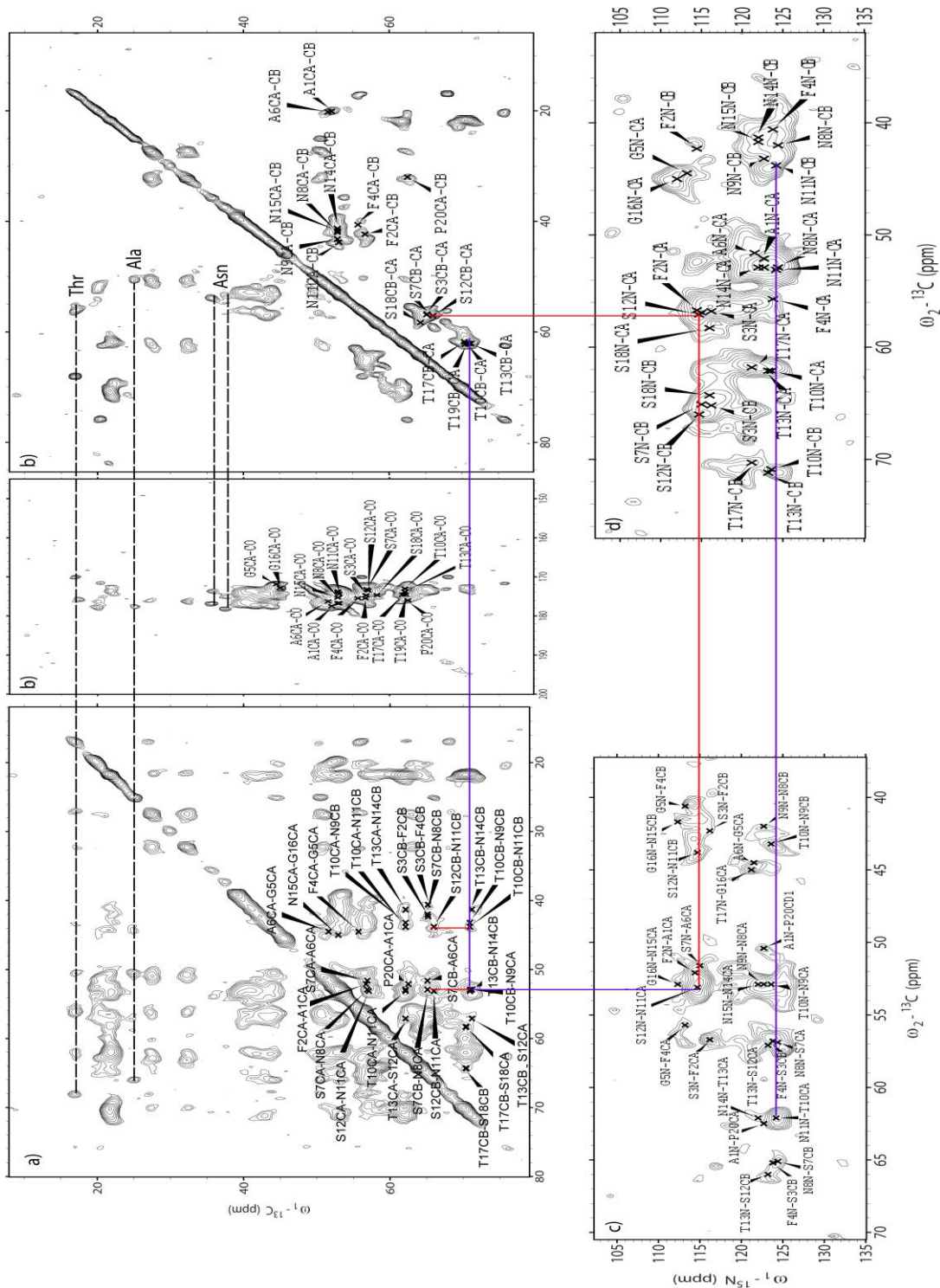


Figure 4: Set of 2D dipolar-based experiments which allowed us to obtain the chemical shift assignments of one strand as listed in Table 2. Identified signals are annotated in the PSD spectrum with short b) and long mixing a), the NCOCX c) and NCACX d) experiment. Assignment strategy is exemplified on the residue combination T10-N11-S12 via red and purple lines for S12-N11 and N11-T10 correlations, respectively. Dashed lines denote outstanding peaks whose resonances could not be included into the assignment of one strand.

spectrum and could be correlated to ^{15}N resonances in the NCACX and NCOCX spectra (see red and purple lines in Figure 4). Only few outstanding peaks could not be included in the assignments within one strand (dashed lines in Figure 4). These signals correspond to spin systems of Asn, Thr and Ala residues. For Ala and Thr sequential correlations are visible (Gly-Ala, Ala-Ser, Thr-Pro). Probably those outstanding signals correspond to residues in the sequence where structural changes take place.

With the assignments at hand (Table 2 lists all assigned chemical shift values), we could compute secondary chemical shifts [11].

number	amino acid	CA	CB	N	CO
1	Ala	52.1	20.1	122.7	177.5
2	Phe	56.7	42.3	114.4	-
3	Ser	56.8	65.2	116.2	-
4	Phe	55.7	40.6	123.8	-
5	Gly	44.5	-	113.2	171.9
6	Ala	51.6	20.2	121.5	176.3
7	Ser	56.9	65.1	115	173.2
8	Asn	52.9	42	124.4	173.7
9	Asn	52.9	43.2	122.7	174
10	Thr	62.1	70.9	123.6	173
11	Asn	53.1	43.8	124.2	174.4
12	Ser	57.1	66	114.7	173.5
13	Thr	62.1	71.2	123.2	173.8
14	Asn	52.9	41.3	122	-
15	Asn	52.9	41.7	122	-
16	Gly	45	-	112	172.8
17	Thr	61.8	70.3	121.2	-
18	Ser	58.3	64.3	116	-
19	Thr	-	70.4	-	-
20	Pro	62.5	32	-	176

Table 2: Chemical shift assignments in ppm of one strand obtained via dipolar-based ssNMR experiments.

The diagrams in Figure 5 show the difference $\Delta C\alpha - \Delta C\beta$ versus the residues of a strand (a) and identified residue types in the INEPT-TOBSY experiment (b). As the Δ refers to the deviation from the random coil value, negative values indicate beta-sheet elements and values around zero are commonly referred to as random coil structure. Apparently, the chemical shift assignments obtained via the dipolar-based experiments contain mainly beta-sheet elements. Particularly, all Asn, Phe and most Ser residues show a strong tendency towards beta-sheet structure. In contrast, the residues which were identified in

the INEPT-TOBSY experiment have random coil conformation. Consequently, the residues belonging to different motional regimes differ in their secondary structures as well.

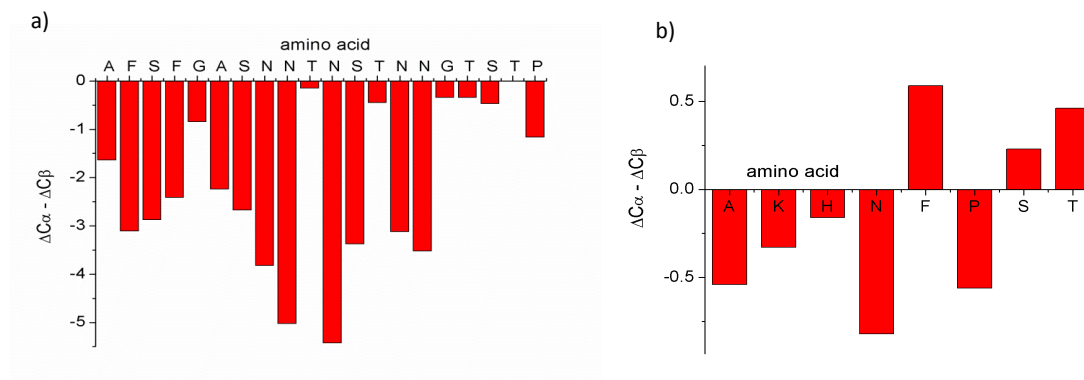


Figure 5: a) and b) illustrate chemical shift difference $\Delta C\alpha - \Delta C\beta$ versus the residues. a) depicts the chemical shift values from dipolar-based experiments and b) from the INEPT-TOBSY. Negative values indicate beta-sheet structure. Values around zero random coil conformation.

Intermolecular contacts and Structural model

To obtain information on intermolecular contacts, we performed a NHHC experiment [12] on a mixed labeled sample which consists of a 1:1 mixture of uniformly ^{13}C and ^{15}N labeled monomers.

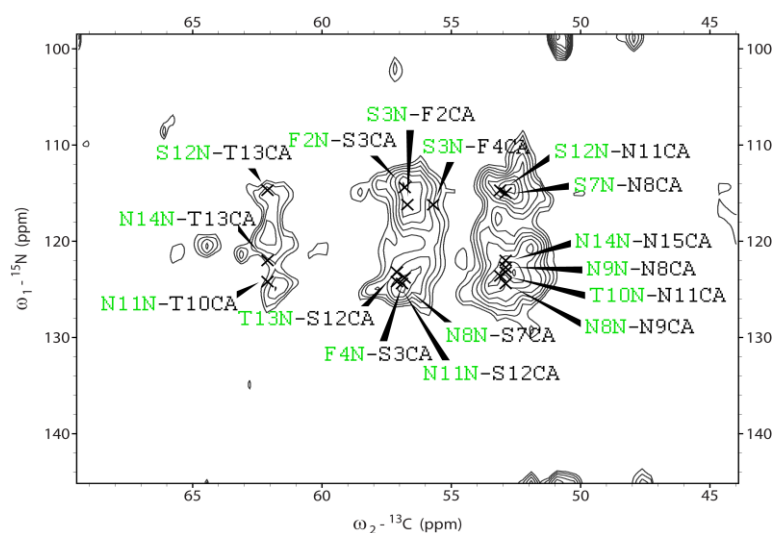


Figure 6: 2D NHHC experiment applied on a $^{13}\text{C}/^{15}\text{N}$ mixed labeled sample. Annotated signals display correlations between the ^{15}N and ^{13}C of different monomers that are indicated in green and black, respectively.

Figure 6 shows the corresponding spectrum with an optimized HH mixing time of 700 μ s. The signals are annotated in different colors for the two monomers in green and black, respectively. Intermolecular contacts were identified on the basis of two monomers consisting of one strand and the chemical shift assignments of Table 2. Signals originating from Ser-Phe correlations and the strong intensity for Asn-Asn correlations can only be explained by a parallel arrangement of beta-sheets belonging to different monomers.

The identified contacts were used as distance restraints in a structure calculation. In addition, dihedral restraints were generated on the basis of the chemical shift assignments.

A structure calculation of 200 structures of two separate strands each 22 amino acids long was performed. Altogether, 34 distance and 84 dihedral restraints were used. Among the 20 lowest energy structures, all strands formed parallel beta-sheets around the Asn-Ser-Thr-Asn region as indicated in Figure 7A.

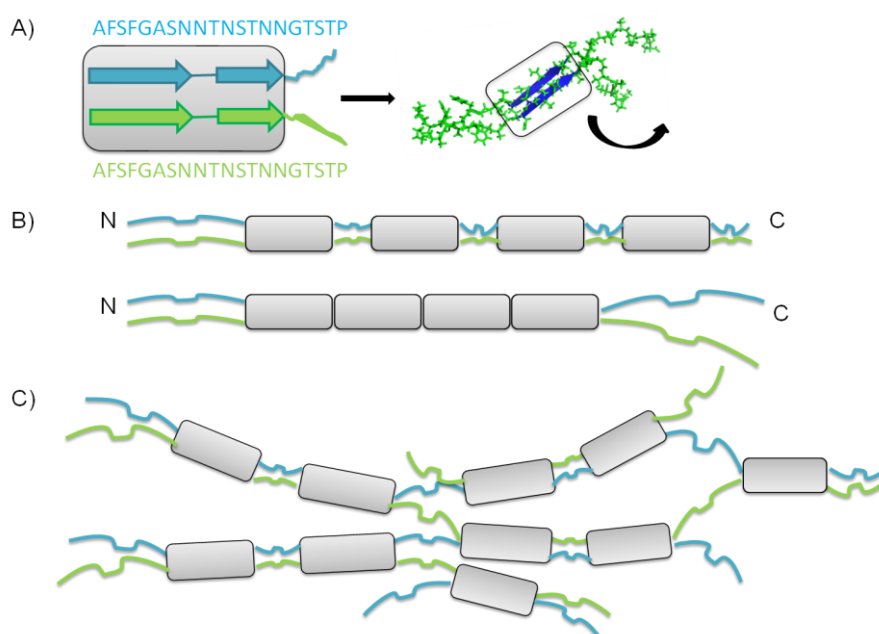


Figure 7: A) Structural model of a protein monomer unit that adopts a parallel beta-sheet arrangement. Distance restraints derived from ssNMR are indicated in red. B) Possible arrangements of the rigid and mobile parts of one repeat unit. In total, there are 16 repeat units per protein monomer. C) Pictorial representation of possible network contacts in the hydrogel.

A few structures showed a second pair of beta-sheets in the Phe region as well. Combined with the results regarding the distribution of rigid and mobile components within the

hydrogel, a first simple structural model was developed. Figure 7B highlights possible arrangements of rigid and mobile repeat units over the protein sequence. Alternating mobile and rigid parts are possible as well as completely connected ones. Finally, Figure 7C suggests possible contacts to form a hydrogel network. Crosslinks are maintained via beta-sheet interactions with different neighboring proteins.

Conclusion

We investigated the dynamical and structural features of a homopolymeric protein hydrogel with sequence similarities to the WT scNsp1 hydrogel by ssNMR measurements. Dipolar-based experiments allowed us to obtain ^{13}C and ^{15}N chemical shift assignments of one strand and revealed beta-sheet elements as its dominating structure. A NHHC experiment on a mixed labeled sample probed intermolecular contacts between parallel arrangements of those beta-sheets to stabilize the hydrogel. Scalar-based experiments showed that around 50 % of the protein is highly mobile and in a random coil conformation. Overall, the homopolymeric and the WT scNsp1 hydrogel are cohesive via beta-sheet structure elements which are essential for forming the hydrogel by hydrophilic interactions.

Furthermore, both hydrogels have significant contributions of highly mobile and rigid regions. These motional regimes were also different in their amino acid constitution in the WT scNsp1 hydrogel e.g. Asn, Gln, Thr and Ser (NQTS spacers) were typical for beta-sheet structure. In case of the homopolymeric variant hydrogel the same residue combinations appeared in both structural forms and motional regimes. Although, the homopolymeric variant with its simplified sequence shows similar transport properties and is stabilized by the same interactions as the WT scNsp1 hydrogel, its repeating units can exhibit different functions.

These findings lead to the general question if the considerable sequence heterogeneity of nucleoporins is required for the functioning of the NPC or if an assembly of nucleoporins with simplified sequences can inherit the same properties. This question could be investigated by future experiments.

Materials and Methods

Preparation of HomoNsp1 Isotope-labeled variants & Hydrogel preparation

Protein and hydrogels were prepared as described previously in [6, 13, 14]. In short, the N-terminally HomoNsp1 domain with a unique C-terminal was expressed in *E. coli* and purified using Nickel-Sepharose and activated Thiopropyl-Sepharose under denaturing conditions. The repeat domain was further purified by reverse phase HPLC, eluted with increasing concentrations of acetonitrile in 0.15% TFA, and lyophilized.

For ssNMR experiments, lyophilized proteins were dissolved at a concentration of 200mg/ml in 200mM Na₂HPO₄. Final preparations displayed a neutral pH and formed hydrogels within 3-4 h at room temperature.

Solid-state NMR experiments

All ssNMR measurements were performed on 3.2 mm triple-resonance (¹H, ¹³C, ¹⁵N) probeheads at 277 K. The proton field strengths for 90° pulses and SPINAL64 [15] decoupling ranged from 74 to 80 kHz and the CP contact times between 0.5 and 0.6 ms. MAS rates were set to 10.92 kHz in (¹³C, ¹³C) and 12 kHz in (¹⁵N, ¹³C) correlation experiments. 2D correlation spectra were obtained using proton-driven spin diffusion with 20 ms and 200 ms DARR mixing [16] for (¹³C, ¹³C) spectra, 60 ms for the NCACX and 60 ms under MIRROR conditions [17] for the NCOCX. In the scalar-based experiments a TOBSY [8] mixing time of 6.6 ms and a GARP [18] decoupling of 10 kHz was used. All experiments were conducted at a static magnetic field strength of 16.4 T except the NCACX which was acquired at 11.7 T. The processing and analysis of NMR data was done using topspin and Sparky [19] software. Dihedral angles were generated based on the chemical shift assignments via PREDITOR [20].

Structure calculation:

The structure calculation was performed using CNS [21]. 200 structures were generated of two separate strands, each with 22 amino acids.

Structural constraints	
Distance restraints from NHHc	34
Dihedral restraints	84

For further investigation only the 20 lowest energy structures were considered.

References

1. Peters, R., *Functionalization of a nanopore: The nuclear pore complex paradigm*. *Biochimica et Biophysica Acta*, 2009. **1793**: p. 1533–1539.
2. Alber, F., et al., *The molecular architecture of the nuclear pore complex*. *Nature*, 2007. **450**: p. 693-701.
3. Ribbeck, K. and D. Görlich, *Kinetic analysis of translocation through nuclear pore complexes* EMBO 2001. **20**: p. 1320-1330.
4. Hülsmann, B.B., A.A. Labohka, and D. Görlich, *The Permeability of Reconstituted Nuclear Pores Provides Direct Evidence for the Selective Phase Model*. *Cell*, 2012. **150**: p. 738-751.
5. Frey, S. and D. Görlich, *A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes*. *Cell*, 2007. **130**: p. 512-523.
6. Ader, C., et al., *Amyloid-like interactions within nucleoporin FG hydrogels*. *PNAS*, 2010. **107**(14): p. 6281-6285.
7. Andronesi, O.C., et al., *Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy*. *Journal of American Chemical Society*, 2005. **127**: p. 12965-12974.
8. Baldus, M. and B.H. Meier, *Total correlation spectroscopy in the solid-state. The use of scalar couplings to determine through-bond connectivity*. *Journal of Magnetic Resonance A*, 1996. **121**: p. 65-69.
9. Castellani, F., et al., *Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy*. *Nature*, 2002. **420**: p. 98-102.
10. Pauli, J., et al., *Backbone and Side-Chain ¹³C and ¹⁵N Signal Assignments of the α -Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla*. *ChemBioChem*, 2002. **2**: p. 272-281.
11. Luca, S., et al., *Secondary chemical shifts in immobilized peptides and membrane proteins: a qualitative basis for structure refinement under magic angle spinning*. *Journal of Biomolecular NMR*, 2001. **20**: p. 325-331.
12. Lange, A., S. Luca, and M. Baldus, *Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids*. *Journal of American Chemical Society*, 2002. **124**(33): p. 9704-9705.
13. Frey, S. and D. Görlich, *FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties*. *EMBO*, 2009. **28**(17): p. 2554-2567.
14. Frey, S., R.P. Richter, and D. Görlich, *FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties*. *Science*, 2006. **314**(5800): p. 815-817.
15. Fung, B.M., A.K. Khitritin, and K. Ermolaev, *An improved broadband decoupling sequence for liquid crystals and solids*. *Journal of Magnetic Resonance*, 2000. **142**(1): p. 97-101.
16. Takegoshi, K., S. Nakamura, and T. Terao, *¹³C-¹H dipolar-assisted rotational resonance in magic-angle spinning NMR*. *Chemical Physics Letters*, 2001. **344**: p. 631-637.
17. Scholz, I., et al., *MIRROR recoupling and its application to spin diffusion under fast magic-angle spinning* *Chemical Physics Letters*, 2008. **460**: p. 278-283.
18. Shaka, A.J., P.B. Barker, and R. Freeman, *Computer-optimized decoupling scheme for wideband applications and low-level operation*. *Journal of Magnetic Resonance*, 1985. **64**: p. 1567-1572.
19. Goddard, T.D. and D. Kneller, *SPARKY 3*. 2006.
20. Berjanskii, M.V., S. Neal, and D.S. Wishart, *PREDITOR: a web server for predicting protein torsion angle restraints*. *Nucleic Acids Research* 2006. **34**(Web Server Issue): p. W63-W69.
21. Brunger, A.T., et al., *Crystallography and NMR System*. 1997, Yale University.

CHAPTER 5

Probing π - π stacking in thermosensitive polymeric micelles with a potential application as drug delivery systems

Related publication: Yang Shi, Mies J. van Steenbergen, Erik A. Teunissen, Luis Novo, Sabine Gradmann, Marc Baldus, Cornelus F. van Nostrum and Wim E. Hennink , *Biomacromolecules*, **14**, 1826-1837 (2013)

Abstract

Self-assembly of thermosensitive amphiphilic block copolymers into micelles complies an important role in drug delivery systems, particularly for hydrophobic drugs. The copolymer mPEG-*b*-p(HPMAm-Nt₁₈-*co*-HPMAm-Lac₈₂) forms micelles above its critical micelle temperature. One- and two-dimensional solid-state NMR experiments detected π - π stacking between the aromatic groups in the micellar core. Furthermore, the copolymer showed high loading capacities for the chemotherapeutic drugs paclitaxel and docetaxel. Altogether, mPEG-*b*-p(HPMAm-Nt₁₈-*co*-HPMAm-Lac₈₂) has a high potential as a new drug delivery system.

Introduction

In the last years, polymeric micelles formed from amphiphilic block copolymers have been extensively investigated as drug delivery systems, particularly for the delivery of hydrophobic drugs [1-3]. Polymeric micelles are characterized by a size normally below 100 nm and a good accommodation for poorly water-soluble drugs. Their hydrophilic corona, mostly consisting of poly(ethylene glycol) (PEG), endows them with a stealth surface likely as a result of a low protein binding to the particles in the blood circulation [4-6]. Among different types of amphiphilic block copolymers, thermosensitive block copolymers have received increased interest for the preparation of micelles [7, 8], hydrogels [9, 10] and to coat liposomes [11]. Thermosensitive diblock copolymers, which have a thermosensitive and a permanently hydrophilic block, form micelles above the critical micelle temperature (CMT) of the thermosensitive block. This yields to polymeric micelles with stealth coronas and hydrophobic cores that can accommodate hydrophobic drugs.

In the following, a thermosensitive copolymer which consists of mPEG, *N*-(2-hydroxypropyl) methacrylamide (HPMAm) modified with a naphthoyl group (HPMAm-Nt) and HPMAm-lactate (HPMAm-Lac) was studied. We performed proton solution- and solid-state (ss) NMR measurements to elucidate the nature of interactions of mPEG-*b*-p(HPMAm-Nt₁₈-*co*-HPMAm-Lac₈₂) during micelle formation. Furthermore, drug content essays demonstrated drug loading capacities with the chemotherapy drugs paclitaxel (PTX) and docetaxel (DTX).

Materials and Methods

Sample preparation

Empty micelles were prepared by a fast heating method as described previously [8]. In short, the polymers were dissolved for 16 hrs at a concentration of 10 mg/mL in 120 mM ammonium acetate buffer (AAB) at pH 5.0 and 0 °C. Next, the polymer solutions were heated in a water bath at 50 °C for 1 min with constant shaking to form micelles. For paclitaxel (PTX) and docetaxel (DTX) loaded micelles, 0.2 mL of the drug solution in ethanol (concentration ranging from 40 to 100 mg/mL) was added to 1.8 mL of an ice cold polymer solution and then immediately heated at 50 °C for 1 min. Subsequently, the micellar dispersions were incubated overnight at room temperature. The free drug was removed by filtration, a method frequently applied to remove non-entrapped drugs from micelles [12-14]. The size of the micelles and the critical micelle temperature (CMT) was determined by DLS. The Z-average hydrodynamic diameters of the empty micelles were 30 nm and between 60 and 70 nm for the drug loaded ones depending on the drug and loading capacity. The CMT was at 10 °C.

Drug content assays

The concentrations of the drugs loaded in the micelles were determined by UPLC analysis using Waters Acquity system consisting of a binary solvent manager, a sample manager and a UV detector. An Acquity® HSS T3 1.8µm column (2×50 mm) was used with a gradient eluent method at a flow rate of 1 mL/min, and a column temperature of 50 °C. Dispersions of drug loaded micelles (0.1 mL) were diluted with 0.9 mL of ACN and subsequently vortexed to destabilize the micelles and dissolve the drug, and then centrifuged at 12.000 g for 10 min. Seven µL of the supernatant was injected and the drug was measured at a wavelength of 227 nm. Samples of the drugs in ACN in a concentration range of 0.2 to 500 µg/mL were used for calibration. Encapsulation efficiency (EE) and loading capacity (LC) were calculated from the UPLC analysis results as follows:

$$EE = \frac{\text{concentration of drug measured}}{\text{concentration of drug added}} \times 100\%$$

$$LC = \frac{\text{concentration of drug measured}}{\text{concentration of (drug measured + polymer added)}} \times 100\%$$

Solution-state NMR experiments

Experiments were conducted at a static magnetic field of 17.6 T. Proton field strengths for 90° pulses ranged between 23 and 27 kHz. Water suppression was achieved by using excitation sculpting with gradients [15].

Solid-state NMR experiments

SsNMR experiments were performed using a Bruker Avance III spectrometer equipped with a 4 mm double resonance (¹H, ¹³C) probehead at 11.7 T static magnetic field. An MAS rate of 1 kHz and an ¹H radio frequency field strength of 66 kHz were used. 2D ¹H-¹H NOESY experiments were conducted using a mixing time of 30 ms. Spectral referencing was done using adamantane.

Experimental results

Solution-state NMR

We performed 1D proton NMR measurements on a micellar suspension prepared in H₂O/D₂O (ratio 20:1) with a polymer concentration of 10 mg/ml above (37 °C) and below (1 °C) the CMT. To minimize water signals in our spectra, we applied a water suppression scheme [15]. Figure 1 illustrates the acquired spectra at both temperatures. Apparently, the aromatic signals disappear completely at 37 °C (red curve). This observation indicates that the size of the formed micelles decreases the molecular tumbling to such an extent that the signals of the aromatic groups broaden. As a result, they cannot be detected anymore by solution-state NMR. As we know the average hydrodynamic diameter of the empty micelles from DLS measurements (30 nm) we can estimate the tumbling of the micelle.

Describing the motion via Brownian rotational diffusion, we can calculate the corresponding rotational correlation time τ_c within the Debye model [16, 17].

$$\tau_c = \frac{4\pi\eta r^3}{3kT}$$

where k denotes the Boltzmann constant and η the viscosity (here water $\eta = 0.7 \times 10^{-4}$ Pas). With $r = 15$ nm radius at $T = 310$ K, we obtain a correlation time of $\tau_c = 0.2$ μ s. As expected this value is rather high compared to typical motional regimes of ns time scale for small inorganic molecules and proteins in solution-state NMR [18].

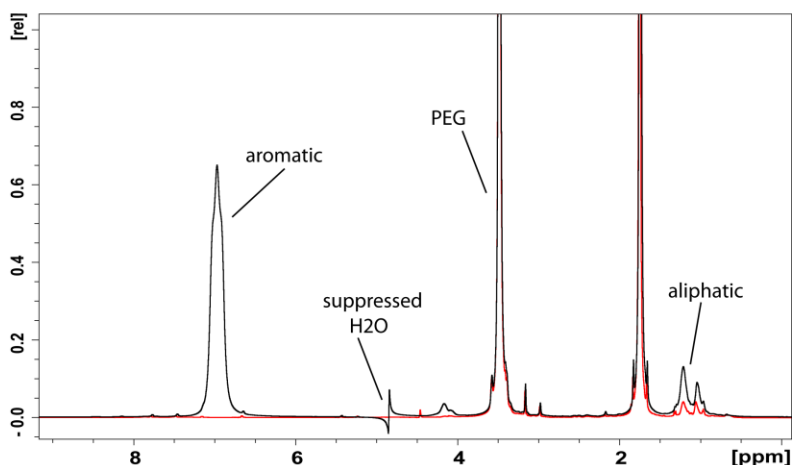


Figure 1: Comparison of 1D solution-state NMR proton spectra obtained at 37 °C (red) and 1 °C (black). Signals of the different parts of the polymer are annotated.

Solid-state NMR

We conducted one and two-dimensional ^1H ssNMR experiments on the micellar suspension prepared in D_2O with a polymer concentration of 10 mg/mL above (at 37 °C) and below (at 1 °C) the CMT. Magic Angle Spinning (MAS, [19]) was employed to increase spectral resolution. In general, such solid-state NMR experiments report on molecular structure via the measurement of isotropic chemical shifts and internuclear interactions, i.e., dipolar spin-spin interactions. The chemical shift of hydrogen atoms is strongly influenced by external electrons providing a means to study packing of π -electron systems. In particular, π - π interactions lead to a shift of aromatic ^1H NMR signals to lower ppm values [20, 21]. In addition, cross peak intensities observed in two-dimensional correlation experiments such as NOESY [22] spectra report on internuclear geometry [20, 23, 24]. Intermolecular interactions such as π - π stacking should hence increase the magnitude and number of cross peak correlations among aromatic protons. Finally, previous solid-state NMR work in semi-solid systems such as hydrogels [25, 26] has shown that a decrease in molecular tumbling leads to an increase in solid-state NMR line width.

Figure 2 compares two 1D ^1H spectra obtained at 37 °C (red) and 1 °C (black) using an MAS rate of 1 kHz. Signals corresponding to the different parts of the copolymer are annotated. Upon increasing the temperature, aromatic signals broadened significantly in line with an overall increase in effective molecular size [25, 26]. Furthermore, aromatic signals shifted towards lower ppm values as expected for π - π interactions. To investigate such interactions by 2D NMR, we acquired 2D NOESY spectra obtained at both

temperatures (Figure 3). In general, cross peaks were more intense at 37 °C (red spectrum) and correlations shifted upfield.

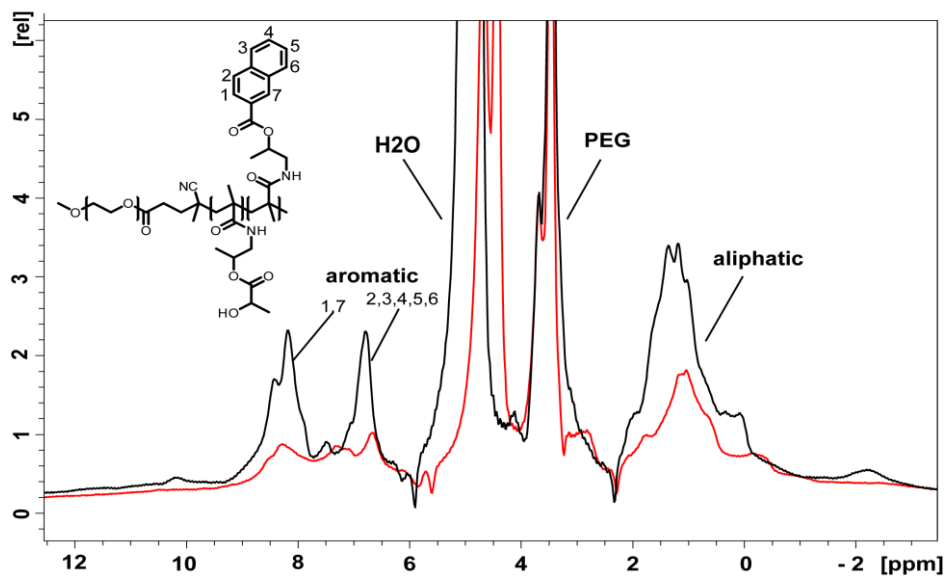


Figure 2: Comparison of 1D solid-state NMR proton spectra obtained under MAS at 37 °C (red) and 1 °C (black). Numbering of the aromatic proton positions is given in the insert.

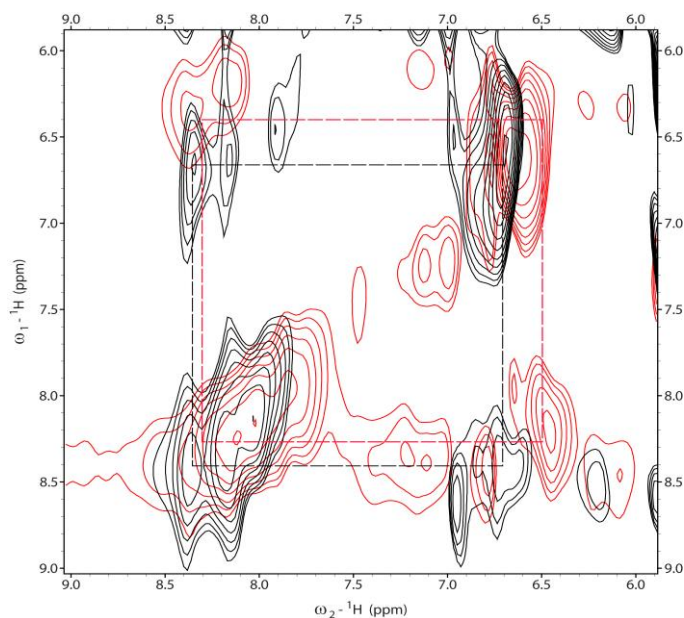


Figure 3: Spectral cutouts of 2D 1H-1H NOESY spectra obtained at 37 °C (red) and 1 °C (black) using a mixing time of 30 ms. Dotted lines connect diagonal and cross-peak correlations.

Although a detailed structural analysis is precluded because cross-peak intensities not only depend on distance but also molecular correlation time and MAS rate [24], the most likely explanation for the observed correlations (indicated by dashed boxes) are proximal intra- or intermolecular nuclear interactions within aromatic rings. Taken together, the NMR results (line width, chemical shifts and 2D cross peak intensities) are consistent with an overall reduction in molecular mobility due to micelle formation that involves π - π stacking [20, 21] among aromatic moieties at higher temperature.

Drug content assay

The drug loading capacity of the micelles was evaluated for chemotherapy drugs PTX and DTX [27] (see Figure 4) that both have aromatic groups and are characterized by a high hydrophobicity (log Ps are 4.7 and 4.1, respectively [28]). When the feed concentrations of drug and polymers were 4 and 7 mg/mL, the loading capacities (LCs) were 27.8 ± 0.3 and 33.5 ± 0.9 % for PTX and DTX, and the encapsulation efficiencies (EEs) were 87.1 ± 3.2 and 64.7 ± 2.1 %, respectively.

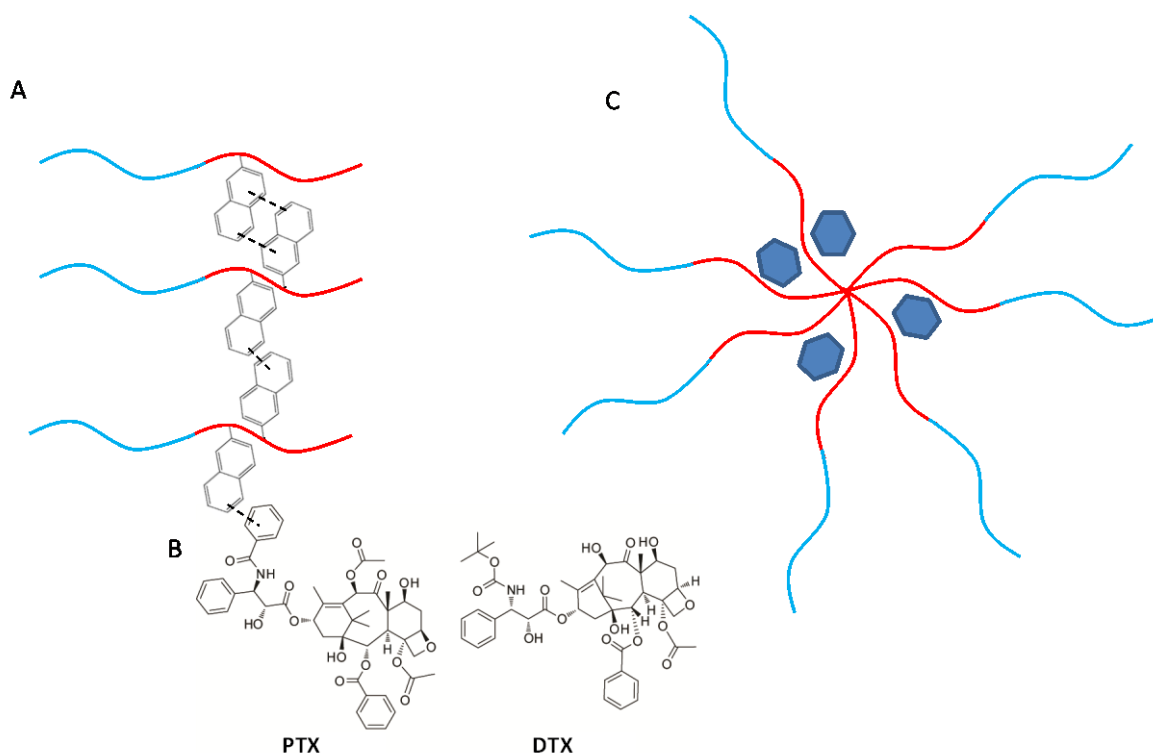


Figure 4: Pictorial representation of a drug loaded micelle and the involved interactions. A depicts a representation of the extended polymer. Hydrophilic and hydrophobic parts are denoted in blue and red, respectively. Dashed lines indicate π - π stacking between the naphthoyl groups and the aromatic group of the drug. In B chemical structures of paclitaxel (PTX) and docetaxel (DTX) are shown. C illustrates micelle formation with incorporation of the drug (blue hexagon).

To the best of our knowledge, polymeric micelles with the highest PTX loading reported in literature are those based on PEG-*b*-p(2-(4-vinylbenzyloxy)-*N,N*-diethylnicotinamide) [29]. For those micelles, the LC of 37.4% for PTX was calculated by weight of drug per weight of polymer. For comparison, when our results are calculated in that way, the LC of the mPEG-*b*-p(HPMAm-Nt₁₈-*co*-HPMAm-Lac₈₂) would be equal to 50.4% for DTX. In other words, we can state that the loading capacity of our micelles is unprecedentedly high. Figure 4 shows a possible incorporation of the drugs into the micelles via π - π stacking among the aromatic groups.

Conclusion

The thermosensitive amphiphilic polymer mPEG-*b*-p(HPMAm-Nt₁₈-*co*-HPMAm-Lac₈₂) self-assembled into micelles above its critical micelle temperature. An investigation by solution-state NMR failed due to the size of the micelles and the resulting slow molecular tumbling. In contrast, ssNMR analysis showed that π - π stacking exists in the micellar core between the aromatic groups. Further drug content assays demonstrated that the poorly water-soluble anticancer drugs paclitaxel and docetaxel were encapsulated into the polymeric micelles with a high loading efficiency. A direct proof of π - π stacking between the aromatic groups of the drug and the polymer will be addressed by future ssNMR experiments.

References

1. Kataoka, K., A. Harada, and Y. Nagasaki, *Block copolymer micelles for drug delivery: design, characterization and biological significance*. *Advanced Drug Delivery Reviews*, 2001. **47**(1): p. 113-131.
2. Torchilin, V., *Micellar nanocarriers: pharmaceutical perspectives*. *Pharmaceutical Research*, 2007. **24**(1): p. 1-16.
3. Talelli, M., et al., *Micelles based on HPMA copolymers*. *Advanced Drug Delivery Reviews*, 2010. **62**(2): p. 231-239.
4. Miller, T., et al., *Analysis of immediate stress mechanisms upon injection of polymeric micelles and related colloidal drug carriers: implications on drug targeting*. *Biomacromolecules*, 2012. **13**(6): p. 1707-1718.
5. Du, J.Z., et al., *Evaluation of polymeric micelles from brush polymer with poly(ϵ -caprolactone)-*b*-poly(ethylene glycol) side chains as drug carrier*. *Biomacromolecules*, 2009. **10**(8): p. 2169-2174.
6. Gaucher, G., R.H. Marchessault, and J.C. Leroux, *Polyester-based micelles and nanoparticles for the parenteral delivery of taxanes*. *Journal of Controlled Release*, 2010. **143**(1): p. 2-12.
7. Wei, H., et al., *Self-assembled, thermosensitive micelles of a star block copolymer based on PMMA and PNIPAAm for controlled drug delivery*. *Biomaterials*, 2007. **28**(1): p. 99-107.
8. Soga, O., C.F. van Nostrum, and W.E. Hennink, *Thermosensitive and biodegradable polymeric micelles with transient stability*. *Journal of Controlled Release*, 2005. **101**: p. 383-385.
9. Censi, R., et al., *In-situ forming hydrogels by simultaneous thermal gelling and Michael addition reaction between methacrylate bearing thermosensitive triblock copolymers and thiolated hyaluronan*. *Macromolecules*, 2010. **43**(13): p. 5771-5778.
10. Hsiue, G.H., et al., *Preparation of controlled release ophthalmic drops, for glaucoma therapy using thermosensitive poly-N-isopropylacrylamide*. *Biomaterials*, 2002. **23**(2): p. 457-462.
11. Paasonen, L., et al., *Temperature-sensitive poly(N-(2-hydroxypropyl)methacrylamide mono/dilactate)-coated liposomes for triggered contents release*. *Bioconjugate Chemistry*, 2007. **18**(6): p. 2131-2136.
12. Huh, K.M., et al., *A new hydrotropic block copolymer micelle system for aqueous solubilization of paclitaxel*. *Journal of Controlled Release*, 2008. **126**(2): p. 122-129.
13. Cho, Y.W., et al., *Hydrotropic agents for study of in vitro paclitaxel release from polymeric micelles*. *Journal of Controlled Release*, 2004. **97**(2): p. 249-257.
14. Jette, K.K., et al., *Preparation and drug loading of poly(ethylene glycol)-block-poly(ϵ -caprolactone) micelles through the evaporation of a cosolvent azeotrope*. *Pharmaceutical Research*, 2004. **21**(7): p. 1184-1191.
15. Hwang, T.-L. and A.J. Shaka, *Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients*. *Journal of Magnetic Resonance A*, 1995. **112**: p. 275-279.
16. Tölle, A. and H. Sillescu, *Rotational Diffusion of Colloid Spheres in Suspension investigated by Deuteron NMR*. *Langmuir*, 1994. **10**: p. 4420-4422.
17. Debye, P., *Polar Molecules*. 1929: Dover, New York.
18. Ishima, R. and D.A. Torchia, *Protein dynamics from NMR*. *Nature Structural Biology*, 2000. **7**(9): p. 740-743.
19. Andrew, E.R., A. Bradbury, and R.G. Eades, *Nuclear magnetic resonance spectra from a crystal rotated at high speed*. *Nature*, 1958. **182**(4650): p. 1659.

20. Schnell, I. and H.W. Spiess, *High-resolution 1H NMR spectroscopy in the solid state: very fast sample rotation and multiple-quantum coherences*. Journal of Magnetic Resonance 2001. **151**(2): p. 153-227.
21. Mafra, L., et al., *Packing interactions in hydrated and anhydrous forms of the antibiotic Ciprofloxacin: a solid-state NMR, X-ray diffraction, and computer simulation study*. Journal of the American Chemical Society, 2011. **134**(1): p. 71-74.
22. Jeener, J., B.H. Meier, and R.R. Ernst, *Investigation of exchange processes by two-dimensional NMR spectroscopy* Journal of Chemical Physics, 1979. **71**: p. 4546-4553.
23. Baldus, M., *Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning*. Progress in Nuclear Magnetic Resonance Spectroscopy 2002. **41**: p. 1-47.
24. Lange, A., et al., *Analysis of proton-proton transfer dynamics in rotating solids and their use for 3D structure determination*. Journal of American Chemical Society 2003. **125**(41): p. 12640-12648.
25. Ader, C., et al., *Amyloid-like interactions within nucleoporin FG hydrogels*. Proceedings of the National Academy of Sciences, 2010. **107**(14): p. 6281-6285.
26. Allix, F., et al., *Evidence of Intercolumnar π - π Stacking Interactions in Amino-Acid-Based Low-Molecular-Weight Organogels*. Langmuir, 2010. **26**(22): p. 16818-16827.
27. Etrych, T., et al., *HPMA copolymer-doxorubicin conjugates: The effects of molecular weight and architecture on biodistribution and in vivo activity*. Journal of Controlled Release, 2012. **164**(3): p. 346-354.
28. Engelkamp, H., S. Middelbeek, and R.J. Nolte, *Self-assembly of disk-shaped molecules to coiled-coil aggregates with tunable helicity* Science, 1999. **284**(5415): p. 785-788.
29. Carstens, M.G., et al., *Small Oligomeric Micelles Based on End Group Modified mPEG-Oligocaprolactone with Monodisperse Hydrophobic Blocks*. Macromolecules, 2007. **40**(1): p. 116-122.

CHAPTER 6

Investigation of Sensory Rhodopsin II and the NpSRII/NpHtrII complex by novel preparations and solid-state NMR methods

Sabine Gradmann, Marc Dittmann, Sarah Ulbrich, Ines Heinrich, Martin Engelhard and
Marc Baldus, Manuscript in preparation

Abstract

New developments in sample preparation and solid-state (ss) NMR methods can provide further insight into the structure and dynamics of biomolecular samples. Here, we investigated the structure of the photoreceptor NpSRII and the NpSRII/NpHtrII complex using three new approaches. Dipolar-based NH experiments applied on a perdeuterated NpSRII sample in a protonated lipid environment under ultra-fast MAS conditions allowed identifying several water-exposed residues and gave preliminary information on protein-lipid contacts. Moreover, cellular preparations of NpSRII were investigated but suffered from limited sensitivity to draw conclusions about possible conformational changes compared to NpSRII in liposomes. Finally, a two-dimensional ^{13}C - ^{13}C direct excitation spectrum performed on the ligated transducer in the NpSRII/NpHtrII complex provided strong hints for a mobile but structured HAMP1 domain.

Introduction

Transmission of signals across cellular membranes is essential for life. Consequently, information regarding the structure and dynamics of proteins that are involved in those processes by responding to external stimuli is highly desirable. A powerful model system is the photoreceptor sensory rhodopsin II (NpSRII) from *Natronomonas pharaonis* that belongs to the family of microbial rhodopsins and triggers a signal transfer cascade homologous to the two-component system of eubacterial chemotaxis which ultimately determines the switch frequency of the flagellar motor.

NpSRII is a seven-helix (A-G) protein with retinal as a cofactor. The retinal is bound to a lysine residue (helix G) via a protonated Schiff base. Upon excitation with light the retinal undergoes an isomerization from an *all-trans* to *13-cis* configuration and induces conformational changes that activate the cognate transducer (NpHtrII). NpHtrII is bound to the SRII by its transmembrane helices (TM1 and TM2). Its remaining part, including two membrane proximal HAMP domains (present in **h**istidine kinases, **a**denylyl cyclases, **m**ethylaccepting chemotaxis proteins and **p**hosphatases), is directed towards the cytoplasmic side [1, 2]. The HAMP domains exist in an equilibrium between a dynamic and compact conformation. The equilibrium between both states can be influenced by different factors (e.g. salt concentration and temperature). Structural analysis of the first HAMP domain using ssNMR led to a model compatible to that published by Hulko et al. [3] comprising a parallel, dimeric, four-helical coiled structure. Further information on the structure of NpSRII, the NpSRII/NpHtrII complex and its different activation states were

obtained by X-ray crystallography [4, 5], EPR spectroscopy [2, 6], solid- [7, 8] and solution-state NMR [9]. Beside a pure structural insight, ssNMR measurements could also probe dynamical features and the water-accessibility of certain residues of the membrane-embedded NpSRII. Furthermore, they revealed an extended receptor-transducer binding interface compared to the available X-ray structure and a mainly well-structured HAMP1 domain within the complex [8]. In spite of the wealth of information obtained so far, a complete understanding of the details of the signal transduction from the receptor to the transducer is still elusive. Consequently, novel NMR methods in combination with innovative membrane protein preparation procedures are desirable to help answer these unresolved questions.

In the following, we present three different approaches that employ ssNMR to gain further insight into the structure of NpSRII in different environments and of the first HAMP domain of the transducer in the membrane-embedded NpSRII/HtrII complex. At first, we investigate perdeuterated NpSRII under ultra-fast MAS conditions. In the second part, we focus on NpSRII in cellular preparations and we finally study NpSRII in complex with the ligated, truncated NpHtrII. All membrane protein samples were prepared in the laboratory of Prof. M. Engelhard (MPI Dortmund).

Materials and Methods

SsNMR experiments

SsNMR experiments were performed on 3.2 and 1.3 mm triple resonance (^1H , ^{13}C , ^{15}N) probeheads at static magnetic fields of 11.7 T, 16.4 T and 18.8 T. Measurements on the 3.2 mm probeheads were conducted using MAS rates of 10 and 12 kHz at a temperature of 263 K. Typical proton field strength for 90° pulses and SPINAL64 [10] decoupling were 76 kHz. ^{13}C magnetization was generated using cross polarization (CP, contact times of 0.4 ms) or direct excitation using 90° ^{13}C pulses at 50 kHz r.f. field strength. (^{13}C , ^{13}C) correlation spectra were obtained using proton-driven spin diffusion schemes with 30 ms mixing or acquiring 2D NOESY spectra with 150 ms mixing.

Measurements on the 1.3 mm probehead were performed employing an MAS rate of 60 kHz at 293 K. Proton and nitrogen field strength for 90° pulses were 135 kHz and 54 kHz, respectively. High power PISSARO decoupling [11] of 126 kHz on protons was applied during the t_1 evolution and low power PISSARO decoupling [12] of 17 kHz on nitrogen was used during acquisition. In both cases, a GARP decoupling [13] of 10 kHz on carbons was

applied. Water suppression was achieved by $(2 \times 10) \times 10$ ms pulses with field strength of 20 kHz. CP contact times in the NH and NHH experiment were 2 ms in the first and 0.6 ms in the second CP step. The ^1H - ^1H mixing was 50 ms. Corresponding pulse sequence is shown in the Supporting Information.

Gene expression and purification of NpHtrI1-94-MXE-His

Transformed *E. coli* BL21(DE3) cells were grown at 37 °C in LB medium containing 100 mg/L ampicillin. At an OD_{600} of 0.8-1.0, 0.5 mM IPTG was added and after an induction period of 4 h, cells were harvested, washed and resuspended in a phosphate buffer (150 mM NaCl, 25 mM disodium hydrogen phosphate, 2 mM EDTA, pH 8). Finally, the cells were broken up in a microfluidizer (Microfluidics Corporation, Newton, USA), membranes were sedimented by centrifugation (45Ti or 70Ti rotor, 45000 rpm, 90 min, 4 °C) and solubilized in buffer A_{mem} (2% DDM, 300 mM NaCl, 50 mM disodium hydrogen phosphate, pH 8) for 16 h at 8 °C. After centrifugation of the solubilized membranes (45000 rpm, 90 min, 4 °C) the supernatant was loaded on a B_{mem} (0.05% DDM, 300 mM NaCl, 50 mM disodium hydrogen phosphate, pH 8) equilibrated column filled with Ni-NTA agarose (Qiagen, Hilden, Germany). The column was washed extensively with buffer C_{mem} (0.05% DDM, 300 mM NaCl, 50 mM disodium hydrogen phosphate, 30 mM imidazole, pH 8) to remove unspecifically bound proteins. Subsequently, the histidine tagged protein was eluted in buffer D_{mem} (0.05% DDM, 300 mM NaCl, 50 mM disodium hydrogen phosphate, 200 mM imidazole, pH 8) and dialysed against a buffer (300 mM NaCl, 50 mM disodium hydrogen phosphate, pH 8). The protein was obtained in a yield of ~ 3.5 mg/ L_{culture} .

Gene expression and purification of GST-TEV-NpHtrIIA95C-157

Transformed *E. coli* BL21(DE3) cells were grown at 37 °C in M9 medium containing 2 g/ L_{culture} ^{13}C -glucose, 1 g/ L_{culture} $^{15}\text{NH}_4\text{Cl}$ (or 2 g/ L_{culture} ^{12}C -glucose, 1 g/ L_{culture} $^{14}\text{NH}_4\text{Cl}$) and 100 mg/L ampicillin. At an OD_{600} of 0.8-1.0, 0.5 mM IPTG was added and after an induction period of 4 h, the cells were harvested, washed and resuspended in a phosphate buffer (150 mM NaCl, 25 mM disodium hydrogen phosphate, 2 mM EDTA, pH 8). Subsequently, the cells were broken up in a microfluidizer (Microfluidics Corporation, Newton, USA) and membrane components were separated by centrifugation (45Ti or 70Ti rotor, 45000 rpm, 90 min, 4 °C). The resulting supernatant was loaded on a GSH Sephadex column, which was equilibrated with PBS buffer (137 mM NaCl, 2,7 mM KCl, 12 mM disodium hydrogen phosphate, pH 7.4). After washing the column extensively with PBS buffer, the protein was eluted with 10 mM glutathione in PBS buffer. The eluate

was concentrated using an Amicon Centrifugal Filter Device (Merck, Darmstadt, Germany) and dialysed against PBS buffer. TEV cleavage of the fusion protein was performed using 1 mg TEV protease per 50 mg GST-TEV-*NpHtrIIA95C-157* with the addition of 1 mM DTT. The proteolytic digestion was monitored using SDS-Page and analytical HPLC (C4 column). After an almost quantitative cleavage, the reaction mixture was centrifuged to remove a white precipitate which was formed during the TEV cleavage. Finally, the supernatant was purified by HPLC using a C4 column (250 x 22 mm, 5 μ m) and a water (buffer A: H₂O, 0.1% TFA)/acetonitrile (buffer B: ACN, 0.08% TFA) gradient (10% B to 70% B in 60 minutes). Fractions were analyzed by ESI-MS and analytical HPLC. Product containing fractions were pooled, lyophilized and stored at -20 °C. The peptide was obtained in a yield of ~ 4 mg/L_{culture}.

Expressed Protein Ligation, Purification and Refolding

The ligation between *NpHtrII1-94-MXE-His* and *NpHtrIIA95C-157* ($[^{12}\text{C}, ^{14}\text{N}]$ or $[^{13}\text{C}, ^{15}\text{N}]$) was performed in ligation buffer (0.05% DDM, 50 mM disodium hydrogen phosphate, 300 mM sodium chloride, 200 mM MESNA, pH 8) with peptide concentrations of ~ 12 μ M at room temperature. The thioester *NpHtrII1-94-MESNA* was generated *in situ*. After two days, the ligation mixture was purified by reversed phase HPLC using a C4 column (250 x 10 mm, 5 μ m) and a formic acid (buffer A: 60% formic acid, 40% water)/isopropanol (buffer B) gradient (5%B to 80%B in 60 minutes). The peak assigned to the ligation product was collected and transferred to a 2 fold excess of TFE. Removal of solvent using a SpeedVac (SpeedVac Plus SC110 A Savant, Savant Instruments inc, New York, USA) and replacement of evaporated solvent by TFE gave a concentrated *NpHtrII1-94-[A95C-157]* ($[^{12}\text{C}, ^{14}\text{N}]$ or $[^{13}\text{C}, ^{15}\text{N}]$) solution in TFE that was diluted with 10 volumes of folding buffer (0.5% DDM, 150 mM NaCl, 25 mM disodium hydrogen phosphate, pH 8). After overnight incubation and dialysis against 100 volumes of dialysis buffer (150 mM NaCl, 25 mM disodium hydrogen phosphate, pH 8) using a membrane with a 10 kDa molecular weight cut-off, the ligation product was centrifuged using a benchtop centrifuge (14000 rpm, 8 °C, 15 min). A small pellet containing *NpHtrII1-94-[A95C-157]* ($[^{12}\text{C}, ^{14}\text{N}]$ or $[^{13}\text{C}, ^{15}\text{N}]$) was observed after centrifugation. The ligation product, refolded in the DDM containing buffer, was obtained in an average yield of 20% after ligation, HPLC purification and refolding.

Reconstitution in proteoliposomes from Halobacterium salinarum

Due to the high DDM concentration in the folding buffer, a direct reconstitution of the semisynthetic *NpHtrII1-157-SRII* complex in proteoliposomes was not possible. Therefore,

a buffer exchange was performed by a Ni-NTA purification of the *NpHtrII1-157-SRII* complex (the used SRII contained a C-terminal His-tag). In order to generate the *NpHtrII-SRII* complex, the semisynthetic *NpHtrII1-157* and SRII were incubated in equal molar quantities overnight. The complex containing solution was loaded on a B_{mem} (0.05 % DDM, 300 mM NaCl, 50 mM disodium hydrogen phosphate, pH 8) equilibrated column filled with Ni-NTA agarose (Qiagen, Hilden, Germany). The column was washed with buffer C_{mem} (0.05% DDM, 300 mM NaCl, 50 mM disodium hydrogen phosphate, 30 mM imidazole, pH 8). Subsequently, the *NpHtrII1-157-SRII* complex was eluted in buffer D_{mem} (0.05% DDM, 300 mM NaCl, 50 mM disodium hydrogen phosphate, 200 mM imidazole, pH 8) and dialysed against buffer B_{mem} . Thereafter, *Halobacterium salinarum* lipids (per 1 mg protein complex one 1 mg lipid was used) and sodium chloride (to reach a final concentration of 1 M) were added and incubated for 30 min at 8 °C. Then, detergent-absorbing Biobeads were added and incubated overnight at 8 °C. After filtration, the reconstituted proteins were pelleted by centrifugation using a benchtop centrifuge (15000 rpm, 4 °C, 15 min), washed five times with a Tris buffer (10 mM Tris-HCl, 150 mM NaCl, pH 8) and were resuspended in the same buffer.

SsNMR results

Perdeuterated Sensory Rhodopsin II (NpSRII)

Recently, perdeuterated proteins with protons reintroduced at exchangeable sites up to different proton levels have gained considerable interest in solid-state (ss) NMR [14-16]. These sparsely protonated samples allow to conduct proton-detected experiments in ssNMR under MAS conditions with a proton resolution that may become comparable to solution-state NMR measurements. Proton linewidths reported range from 0.1-0.2 ppm for membrane-embedded proteins [14] under moderate spinning frequencies (28 kHz, 40% back-exchange, 800 MHz) up to 0.05 ppm for microcrystalline samples [17] at ultra-fast MAS rates (60 kHz, 100% back-exchange, 1 GHz). Together with the general high gyromagnetic ratio of protons and therefore good sensitivity, this approach opens up new perspectives for ssNMR in structural biology. Special tailored experiments allow to obtain proton assignments as well as a large number of long-range ^1H - ^1H contacts in, potentially, short experimental times [18].

Chemical shift assignments and structural restraints are essential for structure determination. Additionally, signals that are present in these spectra may report on the water accessibility of a protein and, in addition, can provide information on the lipid-

protein interface. For example, proton-detected measurements applied on perdeuterated Proteorhodopsin (PR) with different back-exchange levels allowed to assign 30% of the amide protons and revealed information on water-exposed residues [14].

Here, we investigated the use of ^1H -detected ssNMR experiments at ultra-fast MAS conditions using perdeuterated ^{13}C , ^{15}N labeled NpSRII aiming at a protonation level of 100% at water accessible sites. For our ssNMR experiments, NpSRII was reconstituted in synthetic (unlabeled) DMPC lipids. Recent measurements on microcrystalline α -spectrin SH3 have shown that an optimal balance between sensitivity and resolution is obtained for fully back-exchanged samples at 60 kHz and for partial back-exchange (10% up to 30% H_2O) at moderate spinning rates (24 kHz) [17]. Therefore, we performed all our ssNMR experiments under an MAS rate of 60 kHz. In a first step, we acquired a dipolar-based ^{15}N - ^1H experiment. We selected short ^{15}N - ^1H contact times which should lead to the predominant detection of directly bond NH correlations. In Figure 1, experimental results of a two-dimensional ^1H detected NH experiment are shown. Overall, the data are consistent with the spectroscopic analysis of a largely α -helical protein. From well-resolved correlations we estimate residual ^1H and ^{15}N line width of about 0.1 - 0.2 ppm and 1-1.5 ppm, respectively. The latter findings are in reasonable agreement with our previous (^{13}C , ^{15}N) correlation experiments [7] and the proton linewidths are of similar quality as those reported for a partially back-exchanged membrane protein at moderate spinning speeds [14].

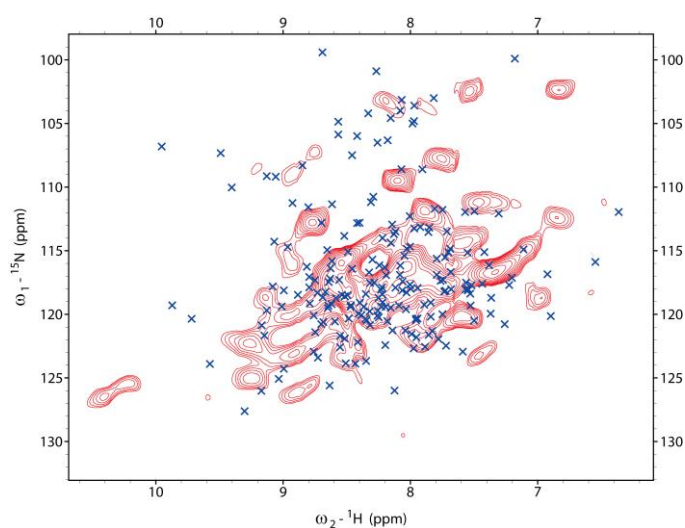


Figure 1: 2D NH experiment applied on perdeuterated NpSRII. Chemical shift predictions in blue are based on solution- and solid-state NMR assignments of NpSRII.

Next, we compared our data to previous NpSR_{II} resonance assignments obtained in liposomes [7] and micelles [9] using FANDAS (see Chapter 2). Former ssNMR experiments on NpSR_{II} have shown that several residues are not visible in dipolar-based experiments [7]. Therefore, we excluded in our further analysis those, potentially mobile, residues from the generated chemical shift predictions. The predictions (blue crosses, Figure 1) in the spectrum are based on solution- [9] and solid-state NMR [7] assignments of NpSR_{II}.

Keeping in mind that ¹H predictions utilized in Figure 1 stem from NMR studies in micelles, we find qualitative agreement between experimental and predicted spectra. In particular, the central, α -helical, region of the spectrum around 120 ppm is well reproduced in the experimental spectrum. A rough estimate of the number of signals and predictions leads to the conclusion that not all predicted correlations are present in the spectrum. This is not surprising as not all residues are expected to be water-accessible.

Figure 2a illustrates water molecules in the crystal structure (PDB-ID: 1H2S) of NpSR_{II}. Obviously, not only loops at the interface are surrounded by water but also several inner helical parts are in contact with water molecules. As a next step, the chemical shift predictions in Figure 1 were limited to residues within a distance of 5 Å from water molecules in the crystal structure. This subset of predictions matches better to the number of acquired signals in the NH experiment (see Figure 2b).

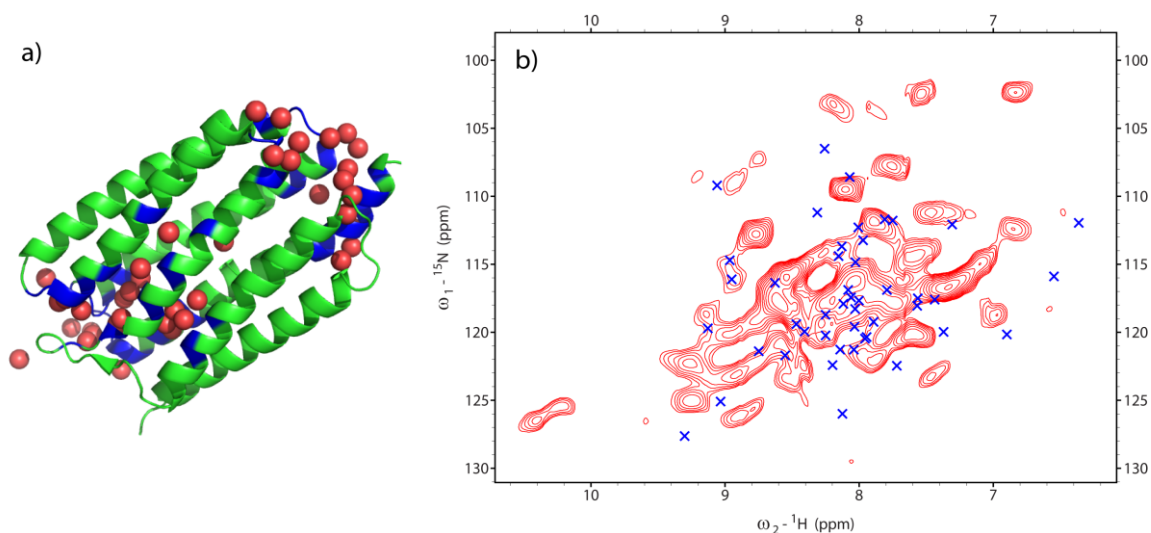


Figure 2: a) Crystal structure of NpSR_{II} (PDB-ID: 1H2S) with water molecules (red). Water-exposed residues within a distance of 5 Å are highlighted in blue. b) Corresponding predictions of the water-exposed residues (blue) are overlaid with a 2D NH experiment.

These remaining residues mainly are in agreement with those detected by water-edited ssNMR experiments on a VLFY reverse-labeled liposome sample before [7]. Only the

region including residues W24 and R27 is devoid of close-by water molecules in the crystal structure in contrast to the ssNMR measurements before. Furthermore, additional residues that are not assigned via ssNMR so far are found to be water-exposed in the crystal structure.

By focusing on several outstanding signals, the reduced number of predictions shows only one prediction in proximity of intensity in some cases and therefore allowed to assign tentatively certain residues. Among the signals which are lacking completely of predictions in case of the reduced prediction data set, S150, I156 and G26 could be identified by the original set of predictions in Figure 1. All identified residues are displayed in Figure 3a and highlighted in green in Figure 3b. In addition, several predictions far-off any intensity were defined as not present at all and consequently not water-exposed (indicated in blue in Figure 3b).

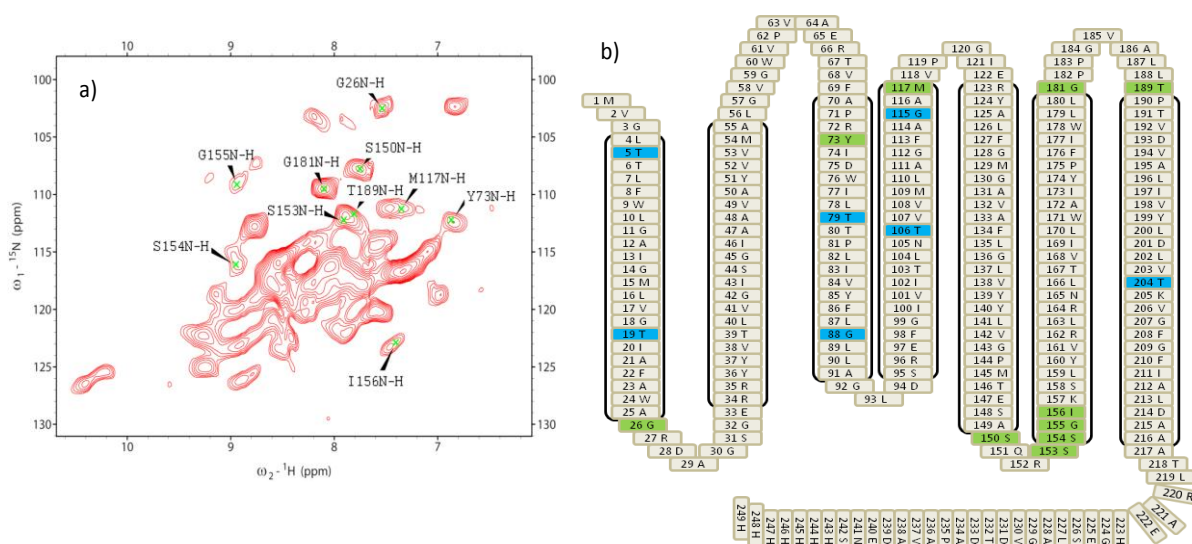


Figure 3: a) 2D NH experiment of NpSR II with several identified residues. b) Sequence scheme of NpSR II. Residues highlighted in green and blue denote water-accessible and non water-accessible ones according to the dipolar-based NH experiment.

In a further experiment, we implemented an additional ^1H - ^1H mixing step of 50 ms. During this mixing time, the magnetization can be transferred further from the amide or other H_2O exchangeable site to surrounding protons. As the lipids were protonated in this sample preparation, we investigated the potential of this experiment to screen for protein-lipid contacts.

Figure 4a shows a 1D ^1H spectrum. The highest intensity stems from water, which was suppressed in the 2D NH before. The broad peak in the region 7-10 ppm results from amide protons. The remaining peaks in Figure 4b can be attributed to lipid signals. In

addition, residual methyl protonation patterns stemming from Leu, Val and Ile for ^1H frequencies between 0-2 ppm might be visible in the spectra [19].

Relating the 1D proton spectrum to the 2D NHH in Figure 4b leads to the following interpretation of the observed signals. The most pronounced peak is the water line. In the region 7-10 ppm the NH spectrum is displayed, however with a reduced sensitivity. The dashed boxes indicate correlations between ^{15}N and not directly attached protons. Intensities in the area highlighted in blue can stem from magnetization transfer to lipids as well as intra-residue transfer to methyl-groups whereas the signals in the region marked in red only correspond to the proton chemical shift of lipids. Published chemical shift values of DMPC obtained by ssNMR [20], attribute this shift to the $\text{R-N}(\text{CH}_3)_3$ in the lipid headgroup. Hence, it is likely that further optimizations including the use of 3D spectroscopy allows to probe receptor – lipid interactions in membranes using proton-detected ssNMR spectroscopy.

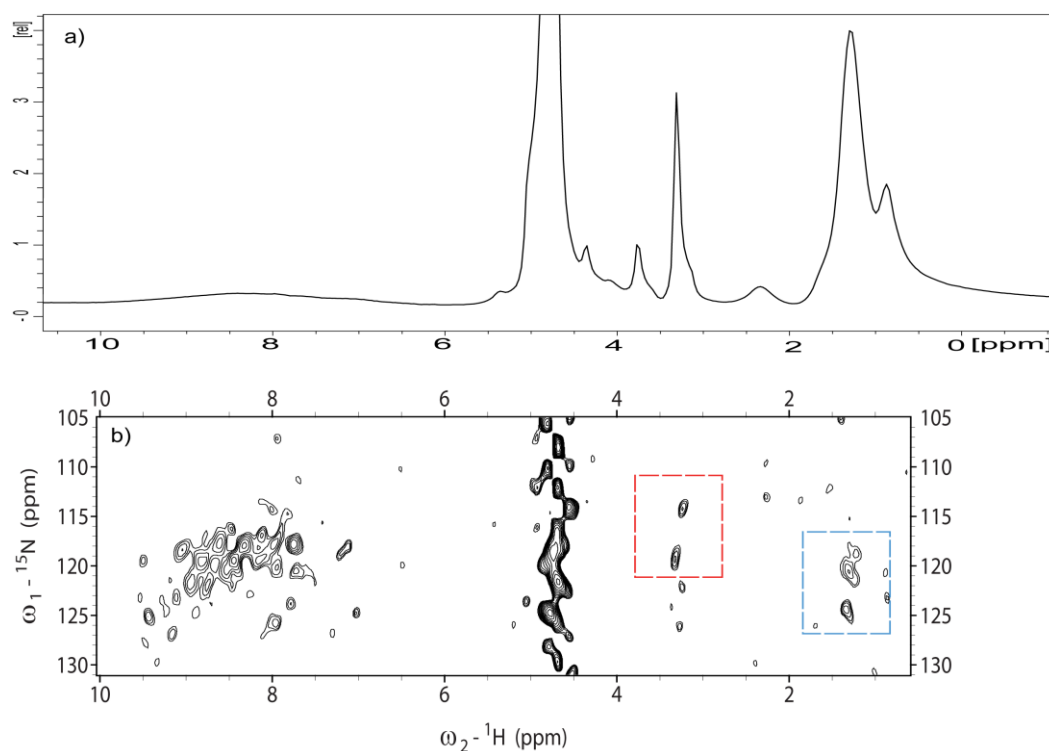


Figure 4: a) 1D proton spectrum and b) 2D NHH spectrum with 50 ms mixing of perdeuterated NpSR11. Dashed boxes indicate correlations between nitrogen and not directly attached protons. Signals stem either from protein-lipid contacts (red and blue box) or further intra-residue transfer (blue box).

Cellular preparations of NpSRII

In addition to using proteoliposome preparations, our group has recently demonstrated that ssNMR studies can be conducted on membrane proteins in a cellular-like environment. As shown for the membrane protein PagL in [21] a set of dipolar- and scalar-based ssNMR experiments allowed to identify and assign several cellular components. Furthermore, it was possible to probe conformational changes of the protein compared to the proteoliposome preparation. In combination with DNP, cellular ssNMR enables further characterization of molecular components [22] as well as monitoring specific interactions e.g. with cell wall polymers (peptidoglycan). Most likely, adjusting of radical concentrations allows to focus on different compartments in the cell as shown in recent studies [23].

Here, we investigated uniformly ^{13}C , ^{15}N labeled NpSRII in a cellular environment. Two different types of preparation were tested. First, the *E. coli* cells were grown in labeled medium from the beginning. In another trial, they were grown in unlabeled medium up to an OD of 0.65 and then were transferred to (^{13}C , ^{15}N) labeled medium with IPTG-induction of NpSRII. As a negative control, a sample of non-induced ^{13}C , ^{15}N labeled cells was used. Figure 5 compares SDS-PAGE gels of an induced and non-induced sample. One recognizes a weak additional band for NpSRII at 26 kDa in the induced cells. Furthermore, a large number of cellular contributions with a peak intensity around 45 kDa is visible.

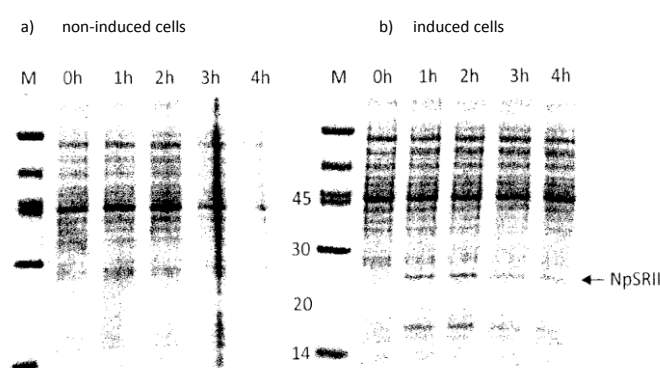


Figure 5: SDS-PAGE gels of a) non-induced and b) induced cells.

For a comparison of both cellular preparations of NpSRII 1D dipolar-based ^{13}C spectra were conducted as shown in Figure 6. As indicated, signals which stem exclusively from cellular components are more intense in the preparation of cells that were grown on labeled medium from the start (blue curve). This is not surprising because the second preparation (red curve) favors the incorporation of isotopes into NpSRII.

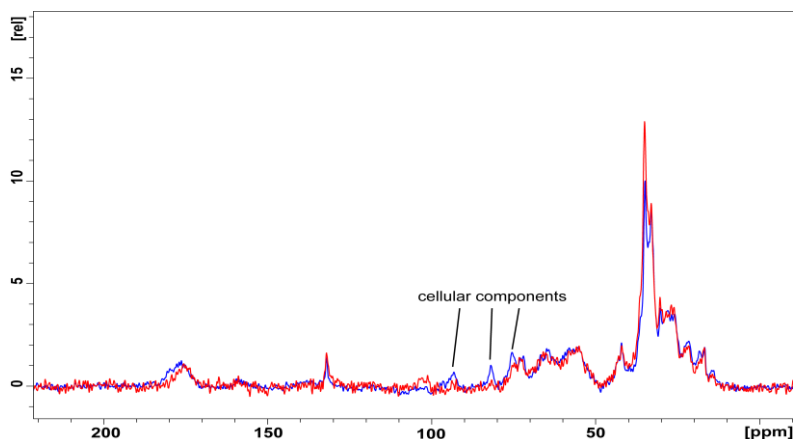


Figure 6: Comparison of 1D dipolar-based ^{13}C spectra of cellular NpSRII. Red and blue curves denote different cellular preparations. Cells were grown in labeled medium from the beginning (blue) or only after reaching a certain OD (red). Signals originating exclusively from cellular components are indicated.

The following experiments were all acquired on the second cellular preparation. Figure 7 shows a 2D (^{13}C , ^{13}C) PDSF spectrum. The overlaid chemical shift predictions stem from crystal structures of proteins which are part of the cell envelope. Contributions from the lipoprotein LPP (PDB-ID: 1EQ7) and the outer membrane protein OmpF (PDB-ID: 2OMF) are shown in black and green, respectively. In principle, all present signals in the aliphatic region can be explained by these cellular compartments. Overall, the signals in the spectrum have α -helical character as shown by the predictions for LPP. Nevertheless, there are certain regions that contain chemical shift values typical for beta-sheets (indicated by dashed boxes in Figure 7). Taking into consideration the main intensities on the gel around 45 kDa, major contributions from the cell would indeed come from the outer membrane protein OmpF and the protein complex cytochrome bo_3 ubiquinol oxidase in the inner membrane [24].

In a further attempt to screen for NpSRII signals in the cellular preparation, a 2D NCA spectrum was acquired. Figure 8 shows an overlay of the NpSRII spectrum (red) with a spectrum obtained on the non-induced, ^{13}C , ^{15}N labeled cells (black). The direct comparison reveals no additional spectral pattern in case of the induced sample which would indicate signals that stem exclusively from NpSRII. Additionally, chemical shift predictions of NpSRII that are based on solution- and solid-state assignments are overlaid with the spectrum. The dashed boxes indicate regions largely devoid of intensity, also in the non-induced case. As the predictions demonstrate the presence of NpSRII signals

there, these areas are critical for future investigations to detect unambiguously signals originating from NpSRII.

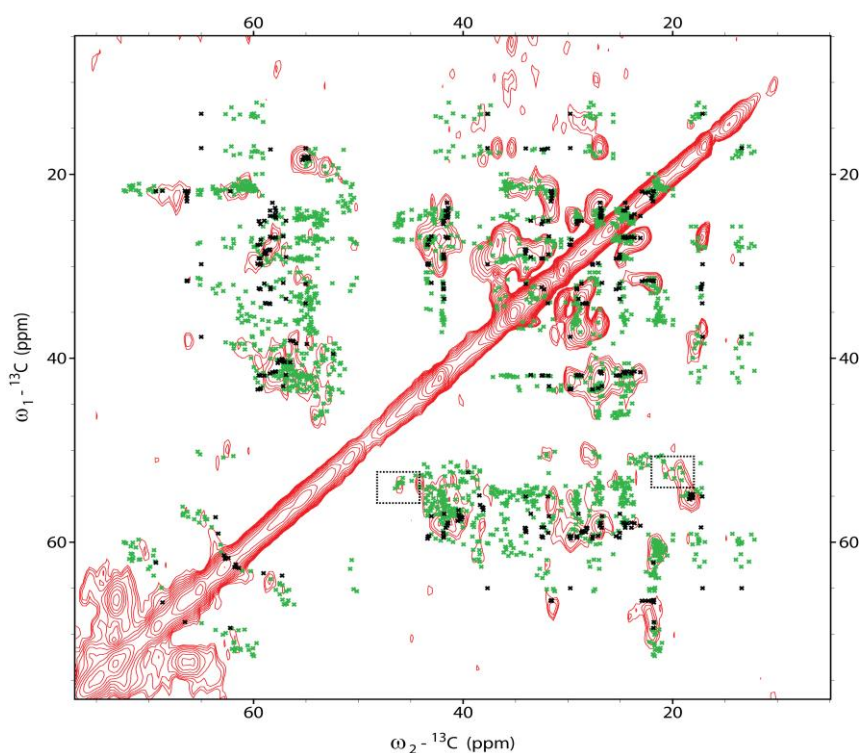


Figure 7: 2D (^{13}C , ^{13}C) PDS spectrum of cellular NpSRII with short mixing times. Chemical shift predictions deduced from the corresponding crystal structures for the lipoprotein LPP (PDB-ID: 1EQ7) and the outer membrane protein OmpF (PDB-ID: 2OMF) are displayed in black and green, respectively. Dashed boxes indicate areas with chemical shift values typical for beta-sheet strands.

Further work, including experiments at lower temperature may help to trace the presence of NpSRII within a cellular setting and reveal possible conformational changes in comparison to preparations in proteoliposomes. At this point, the overall poor sensitivity limits further study. Nevertheless, the current data sets suggest that further high-resolution ssNMR spectra can be obtained to investigate the structure of NpSRII in intact membrane patches or cells.

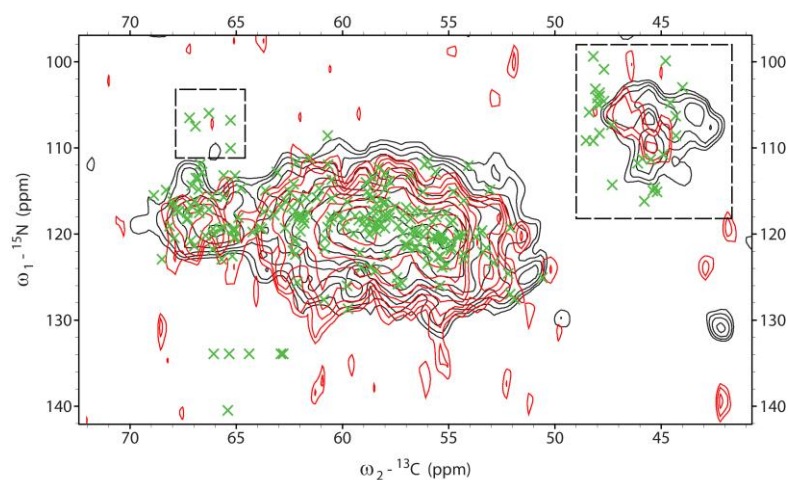


Figure 8: A comparison of 2D NCA spectra of cellular NpSR II (red) and non-induced, labeled cells (black). Chemical shift predictions are based on solution- and solid-state NMR assignments of NpSR II. Dashed boxes indicate regions largely devoid of intensity where signals of NpSR II are expected to be present according to predictions.

Semisynthetic transducer

NMR studies on largely α -helical (membrane) proteins are in general hampered by the limited spectral resolution among backbone ^1H , ^{13}C , ^{15}N resonances. For this reason, advanced labeling procedures such as reverse [7] or pattern [25] labeling are utilized. In addition, segmental isotope labeling technologies using Native Chemical Ligation (NCL), Expressed Protein Ligation (EPL) and Protein Trans-Splicing (PTS) (the methods have been reviewed in [26, 27]) have been used in NMR applications (see, e.g. [28]). A scheme of these methods is depicted in Figure 9.

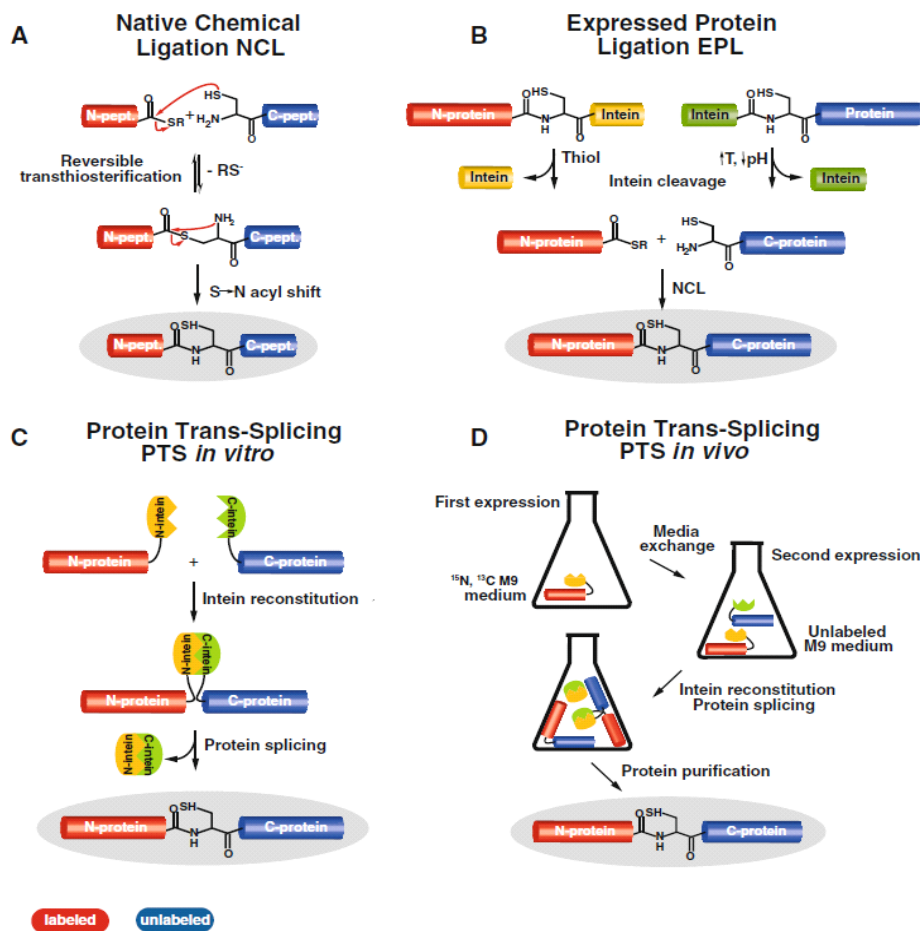


Figure 9: Overview of the methods for segmental isotope labeling of proteins: A, Mechanism of the Native Chemical Ligation. B, Principle of Expressed Protein Ligation. C, Principle of Protein Trans-Splicing by split inteins. D, Schematic representation of the protocol for *in vivo* Protein Trans-Splicing. Reproduced from [28].

Previously, segmental labeling was used to obtain the backbone assignments of the β subunit monomer of F_0F_1 -ATP synthase (52 kDa) by solution-state NMR [29]. Under ssNMR conditions, line widths measured experimentally are generally larger, making the use of segmental isotope labeling technologies an attractive approach. Here, we explored the use of segmental isotope labeling in the context of membrane-embedded NpSR_{II}/NpHtr_{II}¹⁻¹⁵⁷ complex in which transducer which only contains one HAMP domain.

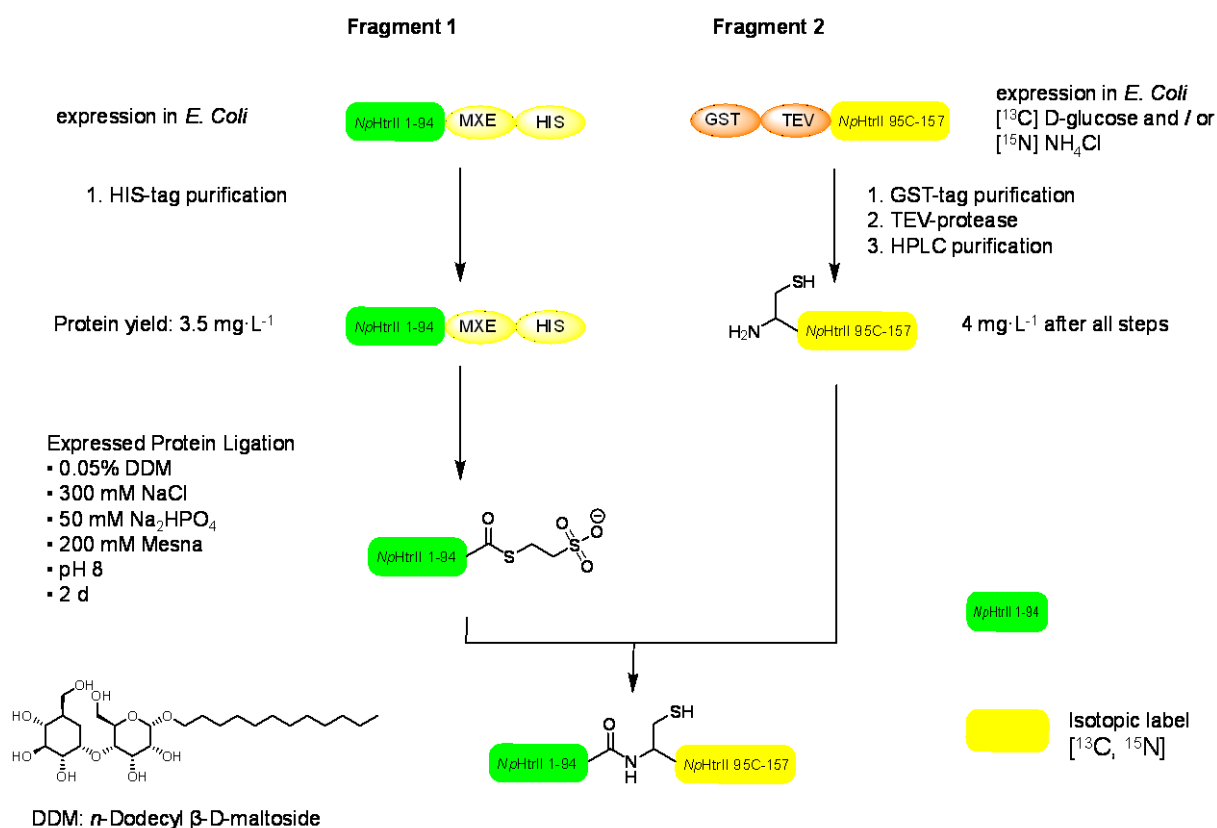


Figure 10: Overview of the Expressed Protein Ligation (EPL) of NpHtrII. (Scheme was kindly provided by Marc Dittmann, MPI Dortmund.)

The sample was prepared in such a manner that only residues 95-157 of the transducer are ¹³C, ¹⁵N labeled (fragment 2) and residues 1-94 (fragment 1) remain unlabeled. This specific labeling was achieved by using EPL (see Figure 10). For the procedure, fragment 1 was expressed as an intein construct (MXE) which is able to generate a thioester. In fragment 2, residue A95 was replaced by a cysteine.

SsNMR experiments that involved scalar- (INEPT) or dipolar-based (CP) magnetization transfer exhibited limited signal-to-noise even at a low sample temperature of 263 K. This may be in part related to low sample concentrations of 1.6 mg transducer. Instead, ssNMR data sets were obtained by using ¹³C direct excitation which is rather insensitive towards protein mobility. Figure 11 illustrates the corresponding 1D and 2D spectra. Although the mixing time in Figure 11b was quite long (150 ms), only a few cross peaks are visible. These intensities could be attributed to Glu, Ser, Ala and Arg residues. Green crosses represent chemical shift predictions of the labeled part of NpHtrII obtained from a homology model based on X-ray and NMR data [8]. Red crosses denote predictions that

are exclusively based on secondary structure information. Here, only random coil formation was taken into account. Whereas the cross peaks in the aliphatic region of the spectrum do not allow to conclude which sort of predictions reflects better the spectral pattern, the carbonyl region points strongly to those predictions based on the homology model. This observation suggests that the labeled part of the transducer is mainly structured in line with previous ssNMR experiments [7].

However, the limited signal-to-noise ratio in the CP based experiment and the necessity of a long mixing time in the 2D spectrum, indicate a rather high mobility. This observation may be reminiscent of an intrinsic molecular plasticity which was proposed in previous EPR studies [2]. In this latter work it has been shown that the HAMP1 domain exists in an equilibrium between two conformations. External conditions like temperature, salt concentration and pH shift this equilibrium. Light favors the more compact conformation (Steinhoff, H. J. and Engelhard, M. unpublished observation). Recent work by Spudich et al. speak in favor of rather well defined HAMP domain structures that undergo synchronized structural changes upon light excitation [30].

Further work will be needed to investigate whether the protein dynamics observed for this construct are due to the absence of the HAMP2 domain (residues 157 – 237). Additional experiments, possibly under low temperature conditions should help to increase signal-to-noise and the number of cross-peak correlations. With these improvements, a more detailed interpretation should be possible.

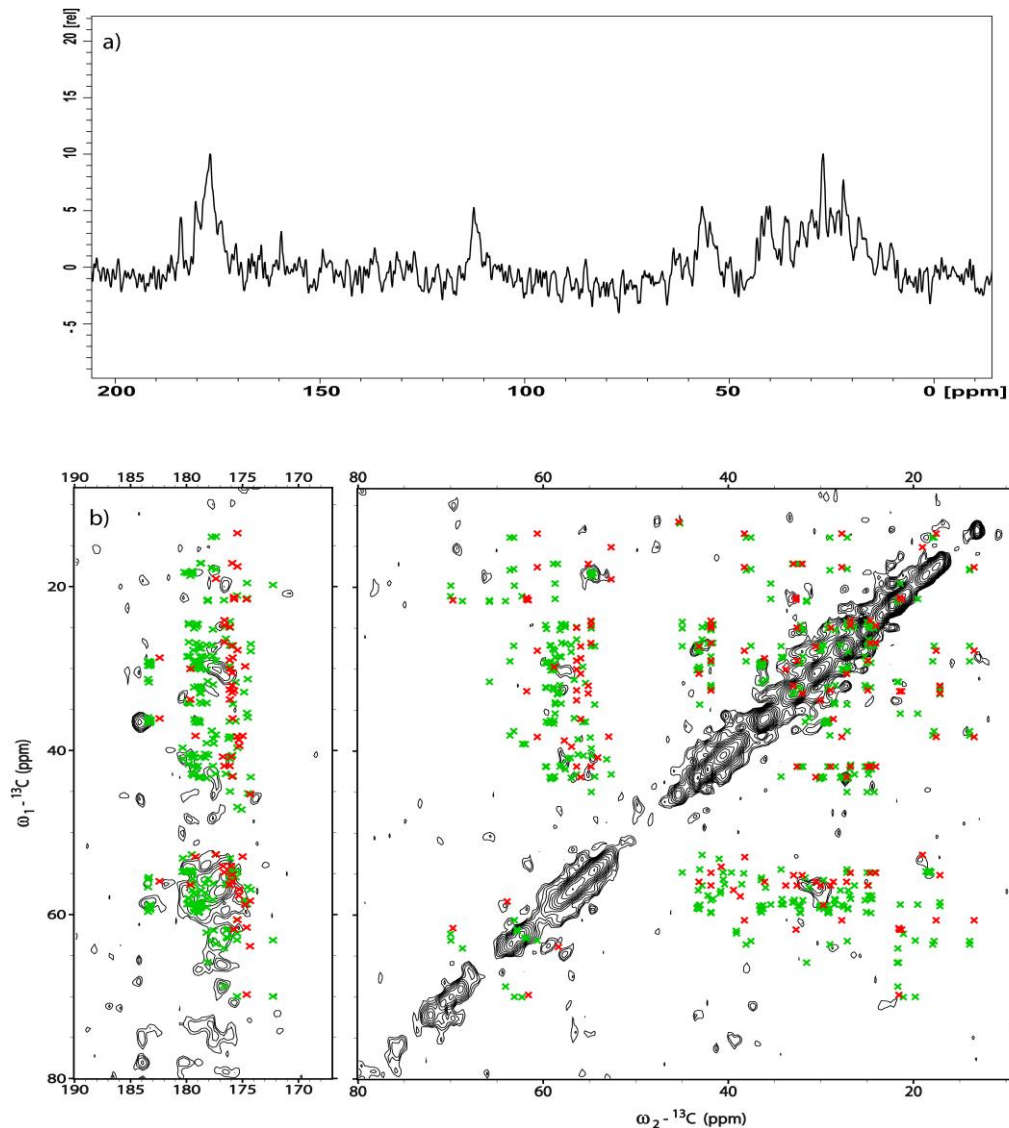


Figure 11: a) 1D ^{13}C spectrum and b) 2D (^{13}C , ^{13}C) NOESY spectrum of the NpSRII/NpHtrII complex. Only the residues 95-157 of the transducer are labeled and therefore included in the predictions. Chemical shift predictions stem from a homology model based on X-ray and NMR data (green crosses) and from secondary structure information according to [31] (random coil conformation, red crosses).

Outlook

Here, we have described three ssNMR-based approaches that are geared towards providing additional information about structure and dynamics of NpSRII and NpSRII/NpHtrII complex in model membranes and its relationship to data in a cellular setting. Clearly, all considered concepts based on a combination of novel preparations

and NMR methods provide promising initial results and suggest additional modifications to facilitate further structural analysis.

In case of the perdeuterated NpSR_{II}, a scalar-based NH spectrum as a complementary experiment would be helpful for a further identification of residues. To overcome the lack of protein signals that stem exclusively from NpSR_{II} in the cellular preparation, DNP measurements at low temperatures which lead to an increase in sensitivity could be useful. Similar DNP experiments should also enhance the structural content and the number of cross peaks in 2D ssNMR spectra obtained on the semisynthetic NpSR_{II}/transducer complex. Also, more advanced correlation schemes, such as dipolar recoupling using RFDR [32] may facilitate polarization transfer within the labeled protein segment. With these improvements, it seems likely that additional information can be obtained regarding the structure and light-activated changes that describe the workings of this membrane protein complex in cellular membranes.

References

1. Wegener, A.A., et al., *Structural insights into the early steps of receptor-transducer signal transfer in archaeal phototaxis*. EMBO, 2001. **20**(19): p. 5312-5319.
2. Doebber, M., et al., *Salt-driven equilibrium between two conformations in the HAMP domain from *Natronomonas pharaonis*: the language of signal transfer?* Journal of Biological Chemistry, 2008. **283**(42): p. 28691-28701.
3. Hulko, M., et al., *The HAMP domain structure implies helix rotation in transmembrane signaling*. Cell, 2006. **126**(5): p. 929-940.
4. Gordeliy, V., et al., *Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex*. Nature, 2002. **419**(6906): p. 484-487.
5. Moukhametzianov, R., et al., *Development of the signal in sensory rhodopsin and its transfer to the cognate transducer*. Nature, 2006. **440**(7080): p. 115-119.
6. Wegener, A.A., et al., *Time-resolved detection of transient movement of helix F in spin-labelled *pharaonis* sensory rhodopsin II*. Journal of Molecular Biology, 2000. **301**(4): p. 881-891.
7. Etkorn, M., et al., *Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy*. Angewandte Chemie International Edition, 2007. **46**(3): p. 459-462.
8. Etkorn, M., et al., *Complex formation and light activation in membrane-embedded sensory rhodopsin II as seen by solid-state NMR spectroscopy*. Structure, 2010. **18**(3): p. 293-300.
9. Gautier, A., et al., *Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy*. Nature Structural & Molecular Biology, 2010. **17**(6): p. 768-775.
10. Fung, B.M., A.K. Khitrin, and K. Ermolaev, *An improved broadband decoupling sequence for liquid crystals and solids*. Journal of Magnetic Resonance, 2000. **142**(1): p. 97-101.
11. Weingarth, M., P. Tekely, and G. Bodenhausen, *Efficient heteronuclear decoupling by quenching rotary resonance in solid-state NMR*. Chemical Physics Letters, 2008. **466**: p. 247-251.
12. Weingarth, M., G. Bodenhausen, and P. Tekely, *Low-power decoupling at high spinning frequencies in high static fields*. Journal of Magnetic Resonance, 2009. **199**(2): p. 238-241.
13. Shaka, A.J., P.B. Barker, and R. Freeman, *Computer-optimized decoupling scheme for wideband applications and low-level operation*. Journal of Magnetic Resonance, 1985. **64**: p. 1567-1572.
14. Ward, M.E., et al., *Proton-detected solid-state NMR reveals intramembrane polar networks in a seven-helical transmembrane protein proteorhodopsin*. Journal of the American Chemical Society, 2011. **133**(43): p. 17434-17443.
15. Linsler, R., U. Fink, and B. Reif, *Assignment of dynamic regions in biological solids enabled by spin-state selective NMR experiments*. Journal of the American Chemical Society, 2010. **132**(26): p. 8891-8893.
16. Reif, B., *Ultra-high resolution in MAS solid-state NMR of perdeuterated proteins: implications for structure and dynamics*. Journal of Magnetic Resonance, 2012. **216**: p. 1-12.
17. Lewandowski, J.R., et al., *Enhanced Resolution and Coherence Lifetimes in the Solid-State NMR Spectroscopy of Perdeuterated Proteins under Ultrafast Magic-Angle Spinning*. Physical Chemistry Letters, 2011. **2**(17): p. 2205-2211.
18. Knight, M.J., et al., *Fast resonance assignment and fold determination of human superoxide dismutase by high-resolution proton-detected solid-state MAS NMR spectroscopy*. Angewandte Chemie International Edition, 2011. **50**(49): p. 11697-11701.

19. Agarwal, V. and B. Reif, *Residual methyl protonation in perdeuterated proteins for multi-dimensional correlation experiments in MAS solid-state NMR spectroscopy*. Journal of Magnetic Resonance, 2008. **194**(1): p. 16-24.
20. Lee, C.W. and R.G. Griffin, *Two-dimensional $^1\text{H}/^{13}\text{C}$ heteronuclear chemical shift correlation spectroscopy of lipid bilayers*. Biophysical Journal, 1989. **55**(2): p. 355-358.
21. Renault, M., et al., *Cellular solid-state nuclear magnetic resonance spectroscopy*. PNAS, 2012. **109**(13): p. 4863-4868.
22. Renault, M., et al., *Solid-state NMR spectroscopy on cellular preparations enhanced by dynamic nuclear polarization*. Angewandte Chemie International Edition, 2012. **51**(12): p. 2998-3001.
23. Takahashi, H., et al., *Solid-State NMR on Bacterial Cells: Selective Cell Wall Signal Enhancement and Resolution Improvement using Dynamic Nuclear Polarization*. Journal of the American Chemical Society, 2013.
24. Stenberg, F., et al., *Protein complexes of the escherichia coli cell envelope*. Journal of Biological Chemistry, 2005. **280**(41): p. 34409-34419.
25. Hong, M. and K. Jakes, *Selective and extensive ^{13}C labeling of a membrane protein for solid-state NMR investigations*. Journal of Biomolecular NMR, 1999. **14**(1): p. 71-74.
26. Vila-Perelló, M., et al., *Streamlined expressed protein ligation using split inteins*. Journal of American Chemical Society, 2013. **135**(1): p. 286-292.
27. Vila-Perelló, M. and T.W. Muir, *Biological applications of protein splicing*. Cell, 2010. **143**(2): p. 191-200.
28. Skrisovska, L., M. Schubert, and F.H. Allain, *Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins*. Journal of Biomolecular NMR, 2010. **46**(1): p. 51-65.
29. Yagi, H., et al., *Conformational change of H⁺-ATPase beta monomer revealed on segmental isotope labeling NMR spectroscopy*. Journal of American Chemical Society, 2004. **126**(50): p. 16632-16638.
30. Wang, J., et al., *HAMP domain signal relay mechanism in a sensory rhodopsin-transducer complex*. Journal of Biological Chemistry, 2012. **287**(25): p. 21316-21325.
31. Wang, Y. and O. Jardetzky, *Probability-based protein secondary structure identification using combined NMR chemical-shift data*. Protein Science, 2002. **11**(4): p. 852-861.
32. Bennett, A.E., et al., *Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange*. Journal of Chemical Physics, 1992. **96**: p. 8624-8627.

CHAPTER 7

Summary and Outlook

Summary and Outlook

Solid-state Nuclear magnetic Resonance (ssNMR) is a versatile spectroscopic method that can be applied to various sample conditions relevant in life and material science and provides atomic insight into molecular structure, dynamics and assembly. This thesis demonstrates the diversity and utility of ssNMR studies in different contexts. It illustrates the potential of ssNMR to determine molecular structure, both de novo or in conjunction with other techniques. In addition, ssNMR is a powerful tool to investigate dynamical features of molecular systems including proteins as well as to monitor self-assembly processes, such as for example regularly occurring in polymers. Finally, all these applications are supported and further enhanced by constant new developments in the field of ssNMR methodology, ranging from dedicated software routines to advanced sample preparation procedures and novel NMR methods.

Chapter 2 introduces the novel software environment and web portal FANDAS (Fast Analysis of multidimensional Nmr DATA Sets) which facilitates ssNMR data analysis of biomolecular spectra and allows to estimate the efficiency of NMR experiments and sample labeling schemes beforehand. The different options of the program were exemplified using applications on the membrane-embedded Sensory Rhodopsin II (NpSRII) and the cyclic nucleotide binding domain (CNBD) of a bacterial cyclic nucleotide-gated channel. Furthermore, it was shown to which extent the incorporation of information from other data sources e.g. X-ray or solution-state NMR supports the assignment process via chemical shift predictions. In this study, a 2D ^{13}C - ^{13}C spin diffusion experiment on NpSRII and a 3D NCOX on CNBD were investigated. Predictions based on 1,3- ^{13}C glycerol labeling on the NpSRII/NpHtrII complex demonstrated the utility of FANDAS data sets to evaluate scrambling effects. A comparison of several reverse labeling strategies of a 2D NCOCA on NpSRII pointed out the possibility to optimize the sample preparation in advance. Furthermore, the example of ^{13}C , ^{15}N labeled NpSRII embedded in ^{13}C labeled lipids showed the potential of the FANDAS software to identify molecular components in large protein complexes or cellular preparations. For the latter case, the inclusion of non-proteinaceous molecular components including lipids or sugar moieties in FANDAS would be desirable and planned in the future.

Nuclear Pore Complexes (NPCs) interconnect the cytoplasmic and intranuclear side of eukaryotic cells and control all nucleocytoplasmic exchange. They consist of approximately 30 nucleoporins (Nups). FG Nups are rich in the amino acids Phenylalanine

and Glycine. Functional studies show that hydrogels formed in vitro by FG Nups exhibit similar transport properties to authentic NPCs. In **Chapter 3** hydrogels of the Nup98 FG nucleoporin and its O-glycosylated version from *Xenopus* were investigated by several ssNMR methods. Dipolar-based (CP, Cross Polarization) experiments, which typically probe rigid parts of the protein, monitored signal intensities from residues across the entire amino-acid sequence in the non-glycosylated case. A fraction of these resonances displayed NMR chemical shifts compatible with the formation of beta-strand folds. After O-GlcNAc modification the CP signal intensity dropped dramatically and indicated a higher mobility of the hydrogel. This was in agreement with the observation that the non-modified Nup98 hydrogel formed a more restrictive permeability barrier than actual NPCs. Consequently, special tailored ssNMR direct excitation experiments that are largely insensitive to protein mobility were applied and revealed a lack of beta-strand folds in the glycosylated case. Instead, clear ssNMR signatures were found that correspond to glycosylated Ser and Thr residues. Comparison to earlier ssNMR results that were obtained on the *S. cerevisiae* (sc) Nsp1p nucleoporin hydrogel suggests that the formation and stabilization of functional hydrogels from FG nucleoporins can be achieved by different mechanisms.

A more detailed analysis of a protein hydrogel is depicted in **Chapter 4**. It focuses on the structural and dynamical characterization of a homopolymeric hydrogel with sequence similarities to the scNsp1p nucleoporin. A set of 2D dipolar-based NMR experiments (PDSO, NCACX, NCOGX) allowed obtaining chemical shift assignments for one monomer unit. The corresponding chemical shift values were typical for beta-strand structures. Intermolecular contacts probed by a 2D NHHX experiment on a mixed labeled sample showed parallel arrangements of adjacent beta-sheets. Dihedral and distance restraints could be deduced from the data and were used for a structure calculation. Additional scalar-based ssNMR experiments identified around 50 % of the protein as highly flexible and unstructured. Therefore, the resulting structural model consisted of monomer units with different structure and dynamical behavior. In combination with the findings in **Chapter 3**, this leads to the conclusion that distinctions in the amino acid constitution of the sequence imply different interactions but also that identical monomers can differ in structure and be involved in different types of interactions.

Besides a structural and dynamical analysis of proteins, NMR also is a tool to trace time-resolved processes including chemical exchange phenomena or self-assembly. In **Chapter 5** studies on the thermosensitive, amphiphilic copolymer mPEG-*b*-p(HPMAm-Nt₁₈-co-HPMAm-Lac₈₂) were performed below and above the critical micelle temperature. 1D and

2D proton ssNMR measurements under low MAS rates were able to detect π - π stacking in the micellar core as a leading interaction. Drug content assays with paclitaxel and docetaxel were successful and showed high loading capacities. Consequently, the investigated polymer has a high potential as a new drug delivery system.

Although standard and established ssNMR experiments have a wide range of applications, a constant development of novel NMR methods enables to access new features and gain more detailed information in systems of increasing complexity. Several new approaches concerning novel preparations as well as NMR methods were discussed in **Chapter 6**. Perdeuterated Sensory Rhodopsin II (NpSR_{II}) in a protonated lipid environment was probed under ultra-fast MAS conditions in terms of water-accessibility and protein-lipid interactions. 2D proton-detected spectra allowed identifying several residues as water exposed and contacts to the lipid headgroup. A further approach was to compare NpSR_{II} in cellular and proteoliposome preparations. So far the findings were not conclusive due to poor sensitivity. A combination with new NMR techniques e.g. DNP should be able to solve this problem in the future. A further application of ssNMR measurements on the ligated and truncated transducer in the NpSR_{II}/NpHtr_{II} complex provided strong hints to a mobile but mainly structured first HAMP domain region. These findings provide a first indication about the structural role of the HAMP1 domain in light-mediated signal transduction in the NpSR_{II}/NpHtr_{II} complex.

Samenvatting en Vooruitzichten

Vaste stof NMR (ssNMR: *solid-state Nuclear Magnetic Resonance*) is een veelzijdige spectroscopische techniek die toegepast kan worden op verschillende relevante monsters uit de levens- en materiaalwetenschappen om met een atomistisch detail inzicht te krijgen in de moleculaire structuur, dynamica en assemblage. Dit proefschrift laat onder verschillende omstandigheden de diversiteit en het nut van ssNMR studies zien. Het toont aan dat ssNMR de potentie heeft moleculaire structuren op te helderen, zowel *de novo* als in combinatie met andere technieken. Daarnaast is ssNMR een krachtige techniek als het gaat om het in kaart brengen van het dynamische karakter van moleculaire systemen zoals eiwitten, als ook het waarnemen van zelfassemblage processen die bijvoorbeeld vaak plaatsvinden bij polymeren. Tenslotte worden deze toepassingen voortdurend ondersteund en versterkt door nieuwe ontwikkelingen op het gebied van ssNMR, variërend van speciale software routines en nieuwe NMR methoden tot geavanceerde procedures voor het bereiden van monsters.

Hoofdstuk 2 beschrijft het nieuwe programma FANDAS (Fast Analysis of multidimensional Nmr DATA Sets) en de bijbehorende webportaal, waarmee de analyse van biomoleculaire ssNMR spectra vergemakkelijkt wordt en het mogelijk is op voorhand efficiënte isotooplabele strategieën en bijbehorende NMR experimenten uit te werken. De verschillende opties van het programma worden beschreven aan de hand van NMR data aan het membraan eiwit *Sensory Rhodopsin II* (NpSRII) en het cyclisch nucleotide bindend domein (CNBD) van een bacterieel cyclisch nucleotide gestuurd kanaal. Daarnaast wordt laten zien in hoeverre het toevoegen van informatie van andere origine, zoals van röntgendiffractie of vloeistof NMR experimenten, de toekenning van de chemische verschuivingen bevordert. Zowel een 2D ^{13}C - ^{13}C spin-diffusie experiment aan NpSRII als ook een 3D NCOCA aan CNBD worden in deze studie uitgewerkt. Vervolgens wordt laten zien dat FANDAS voorspellingen zeer nuttig zijn om scrambling effecten te evalueren, bijvoorbeeld in het geval van het 1,3- ^{13}C glycerol gelabeld NpSRII/NpHtrII complex. Het vergelijken van verschillende omgekeerde isotooplabele strategieën toegepast op een 2D NCOCA experiment aan NpSRII illustreert de mogelijkheid om op voorhand de monster voorbereiding te optimaliseren. Daarnaast laat het voorbeeld van ^{13}C - ^{15}N gelabeld NpSRII in ^{13}C gelabelde lipiden zien dat FANDAS gebruikt kan worden om moleculaire componenten in grote eiwit complexen of cellulaire bereidingen te identificeren. In het laatste geval is het wenselijk dat ook niet-eiwit moleculaire componenten, zoals lipiden of suikergroepen in FANDAS worden opgenomen, wat reeds gepland is voor de toekomst.

Nucleaire Pore Complexen (NPCs) verbinden het cytoplasma en de kern in eukaryote cellen met elkaar en bewaken alle uitwisselingsprocessen tussen kern en cytoplasma. Ze bestaan uit ongeveer 30 nucleoporins (Nups). FG Nups zijn rijk aan de aminozuren Fenylalanine en Glycine. *In vitro* gemaakte hydrogelen van deze FG Nups vertonen dezelfde transport eigenschappen als authentieke NPCs. In **Hoofdstuk 3** worden de *in vitro* hydrogelen van *Xenopus* Nup98 FG nucleoporin en de O-geglycolyseerde versie bestudeerd met behulp van verschillende ssNMR experimenten. Dipolaire (CP, Cross Polarization) experimenten, die voornamelijk de rigide delen van een eiwit laten zien, vertonen signalen van alle aminozuren in het geval van het niet geglycolyseerde eiwit. Een aantal van deze resonanties hebben chemische verschuivingen die overeen komen met een beta-strand vouwing. Na O-GlcNAc modificatie was de signaalintensiteit echter dramatisch lager, wat overeenkomt met een grotere mobiliteit van de hydrogel. Dit is in lijn met de observatie dat een niet gemodificeerde Nup98 hydrogel een lagere permeabiliteit vertoont dan de eigenlijke NPCs. De specifiek afgestemde directe excitatie ssNMR experimenten die vervolgens werden toegepast en grotendeels ongevoelig zijn voor eiwit dynamica, lieten zien dat in het geval van glycosylering de beta-strand vouwing afwezig was. In plaats daarvan werden duidelijke ssNMR kenmerken waargenomen van geglycosyleerde Ser en Thr residuen. De vergelijking met eerdere ssNMR resultaten aan de *S. cerevisiae* (sc) Nsp1p nucleoporin hydrogel suggereert dat de vorm en stabiliteit van functionele FG nucleoporin hydrogelen via verschillende mechanismes tot stand kunnen komen.

Een meer gedetailleerde analyse van een eiwit hydrogel wordt beschreven in **Hoofdstuk 4**. Er wordt in gegaan op de structurele en dynamische karakterisering van een homopolymere hydrogel die qua aminozuur volgorde gelijkenissen vertoont met het scNsp1p nucleoporin. Met behulp van een aantal 2D dipolaire NMR experimenten (PDSO, NCACX, NCOGX) konden de chemische verschuivingen van één monomere eenheid worden toegekend. De gevonden waarden voor de chemische verschuivingen kwamen duidelijk overeen met die van een beta-strand structuur. Doormiddel van een 2D NHHC experiment aan een gemixt isotoop gelabeld monster kon de antiparallelle organisatie van twee aangrenzende beta-strands bepaald worden. De dihedrale- en afstands-restraints die met de data bepaald konden worden werden gebruikt om een structuur te berekenen. Aanvullende ssNMR experimenten gebaseerd op scalaire J-koppeling onthulde dat ongeveer 50 % van het eiwit zeer flexibel en ongestructureerd is. Daarom bestaat het resulterende model voor de structuur uit monomere eenheden met een variërende structuur en dynamisch gedrag. Samen met de resultaten uit **Hoofdstuk 3** kan worden geconcludeerd dat verschillen in de aminozuur volgorde zorgen voor andere

interacties, maar dat ook identieke monomeren kunnen verschillen van structuur en dus ook andere interacties maken.

Naast de analyse van eiwit structuren en hun dynamica vertegenwoordigt NMR ook een methode waarmee tijdsafhankelijke processen, zoals chemische uitwisseling of zelf-assemblage processen, gevolgd kunnen worden. **Hoofdstuk 5** beschrijft experimenten aan de temperatuur gevoelige en amfifiele co-polymeer mPEG-*b*-p(HPMAM-Nt₁₈-CO-HPMAM-Lac₈₂), die zowel onder en boven de kritische micelle temperatuur werden uitgevoerd. Met behulp van 1D en 2D proton ssNMR metingen bij een lage MAS frequentie kon π - π stapeling als belangrijkste interactie in de micelle kern worden geïdentificeerd. Een geneesmiddel assay met paclitaxel en docetaxel was succesvol en een hoog ladingsvermogen kon worden aangetoond. Hierdoor is de bestudeerde polymeer een potentiële kandidaat voor het gebruik als nieuw geneesmiddel afgiftesysteem.

Ook al hebben de standaard en gebruikelijk ssNMR experimenten een breed scala aan toepassingen, worden er continu nieuwe NMR methodes ontwikkeld die nieuwe informatie verschaffen en een meer gedetailleerde inzicht geven in systemen van toenemende complexiteit. Verschillende nieuwe manieren om monsters te maken als ook nieuwe NMR methoden worden beschreven in **Hoofdstuk 6**. Interacties met water en/of lipiden werden bestudeerd in geperdeutereerd Sensory Rhodopsin II (NpSRII) in een geprotoneerde lipide omgeving bij een ultrahoge MAS conditie. Door middel van 2D proton-gedeteteerde experimenten konden verschillende residuen in contact met water en de lipide hoofdgroep worden geïdentificeerd. Daarnaast werden de NMR spectra van NpSRII in een cellulaire en proteoliposoom omgeving vergeleken. Door de lage gevoeligheid konden tot dusver hierover nog geen conclusies getrokken worden. Dit probleem zou in de nabije toekomst opgelost moeten kunnen worden door nieuwe NMR methoden zoals DNP. Een andere toepassing van ssNMR metingen aan een geligeerde en getrunceerde transducer in het NpSRII/NpHtrII complex gaf duidelijk aan dat het eerste HAMP domein flexibel maar wel voornamelijk gestructureerd is. Deze bevindingen geven een eerste inzicht in de structurele rol van het HAMP1 domein in licht gemedieerde signaal transductie in het NpSRII/NpHtrII complex.

Supporting Information

Chapter 2

Supporting Table 1: FANDAS input nomenclature

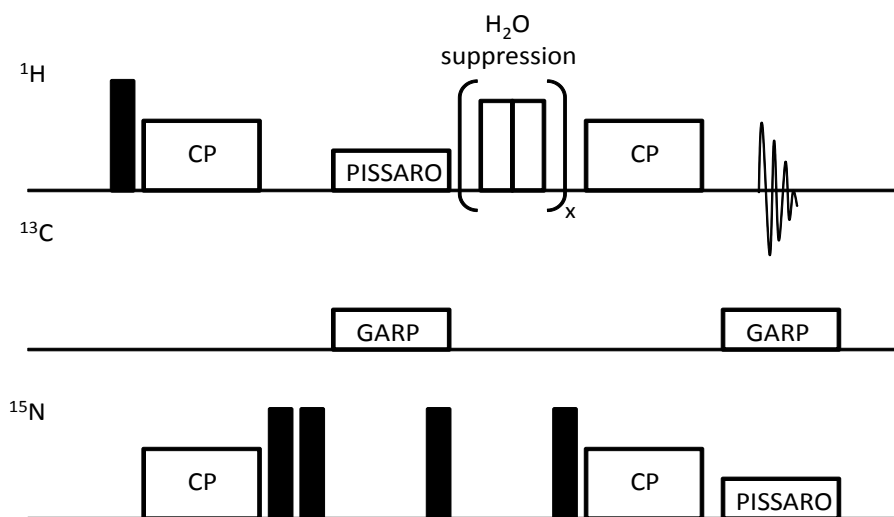
Entry	BMRB format	Deviations from BMRB format
cs[X] _{i,1}	H	HN for all amino acids
cs[X] _{i,2}	HA	HA= HA2=HA3: G
cs[X] _{i,3}	HB	HB=HB2= HB3: R,D,N,C,E,Q,H,L,K,M,F,P,S,W,Y
cs[X] _{i,4}	HG	HG=HG12=HG13:I HG=HG1: V HG=HG2=HG3: R,E,Q,K,M,P
cs[X] _{i,5}	HG2	
cs[X] _{i,6}	HD	HD = HD1: H,I,L,F,W,Y HD = HD2=HD3: R,K,P
cs[X] _{i,7}	HD2	HD2 = HD21=HD22: N
cs[X] _{i,8}	HE	HE=HE1: H,F,W,Y HE=HE2,HE3: K
cs[X] _{i,9}	HE2	HE2=HE21,HE22: Q
cs[X] _{i,10}	HE3	
cs[X] _{i,11}	HZ	
cs[X] _{i,12}	HZ2	
cs[X] _{i,13}	HZ3	
cs[X] _{i,14}	HH	HH=HH2: W, HH=HH11=HH12: R
cs[X] _{i,15}	HH2	HH2=HH21=HH22: R
cs[X] _{i,16}	N	nitrogen (backbone) chemical shift
cs[X] _{i,17}	C	CO (backbone) chemical shift
cs[X] _{i,18}	CA	
cs[X] _{i,19}	CB	
cs[X] _{i,20}	CG1	CG1=CG: R,D,N,Q,E,H,L,K,M,F,P,W,Y
cs[X] _{i,21}	CG2	
cs[X] _{i,22}	CD1	CD1=CD: R,E,Q,K,P
cs[X] _{i,23}	CD2	
cs[X] _{i,24}	CE1	CE1=CE: K,M
cs[X] _{i,25}	CE2	
cs[X] _{i,26}	CE3	
cs[X] _{i,27}	CZ1	CZ1=CZ: Y,F,R
cs[X] _{i,28}	CZ2	
cs[X] _{i,29}	CZ3	
cs[X] _{i,30}	CH	CH=CH2: W

Supporting Table 2: Currently supported 2D and 3D experiments

Dimension	Nuclei	(Coherence order) - Transfer type step 1 (& transfer type step 2) - (abbreviation)
2D	$^{13}\text{C} - ^{13}\text{C}$	intra-residue dipolar/scalar – (CC) inter-residue dipolar/scalar – (CC) distance encoded - (CC) HH distance encoded – (CHHC)
2D	$^{13}\text{C} - ^{13}\text{C}$	(DQ) –one-bond dipolar – (DQSQ CC)
2D	$^{15}\text{N} - ^{13}\text{C}$	one-bond dipolar – (NCA) one-bond dipolar – (NCO) intra-residue dipolar – (NCACX) inter-residue dipolar- (NCOCA/CB, NCOCX) HH distance encoded – (NHHC)
2D	$^1\text{H} - ^{15}\text{N}$	One-bond dipolar – (HN)
2D	$^1\text{H} - ^{13}\text{C}$	One-bond dipolar – (HC)
2D	$^{13}\text{C} - ^1\text{H}$	One-bond-dipolar – (CH)
2D	$^1\text{H} - ^1\text{H}$	intra-residue or distance encoded – (HH)
2D	$^{13}\text{C} - ^1\text{H}$	H-H distance encoded – (CHH)
2D	$^1\text{H} - ^{13}\text{C}$	H-H distance encoded - (HHC)
3D	$^1\text{H} - ^1\text{H} - ^{13}\text{C}$	H-H distance encoded & one-bond dipolar - (HHC)
3D	$^{15}\text{N} - ^1\text{H} - ^1\text{H}$	One-bond dipolar & H-H distance encoded – (NHH)
3D	$^{13}\text{C} - ^1\text{H} - ^1\text{H}$	One-bond dipolar & H-H distance encoded – (CHH)
3D	$^{15}\text{N} - ^{13}\text{C} - ^{13}\text{C}$	One-bond dipolar & intra/inter residue dipolar or distance encoded CC – (NCACX, NCOCX)
3D	$^{13}\text{C} - ^{13}\text{C} - ^{13}\text{C}$	Dipolar/scalar intra residue & dipolar/scalar inter residue – (CCC)
3D	$^{13}\text{C} - ^{13}\text{C} - ^{13}\text{C}$	(DQ) One-bond dipolar & dipolar intra/inter residue SQQ – (DQC-SQ(C,C))

Chapter 6

Supporting Figure 1: Pulse sequence used for the NH experiment. Black filled rectangles represent 90° pulses.



Acknowledgements

Here it goes, four years have passed and the encounters with many people crossed my way. Each of them contributed in their own way to this thesis.

First of all, I thank my supervisor Marc Baldus for the interesting projects he entrusted me with, his support and for giving me the opportunity to be a fellow of the Marie Curie ITN SBMP.

I am indebted to all the fruitful collaborations that I could benefit from: The group of Martin Engelhard (Max-Planck-Institute of Molecular Physiology, Germany) for the preparation of the NpSRII samples with particular thanks to Marc Dittmann, Sarah Ulbrich and Ines Heinrich; Aksana Labohka, Steffen Frey and Dirk Görlich (Max-Planck-Institute for Biophysical Chemistry, Germany) for the preparation and functional tests of the hydrogels and Shi Yang, Wim Hennink and Cornelus van Nostrum (Utrecht University, the Netherlands) for their work on the polymeric micelles.

Further, I thank Alexandre Bonvin, Rolf Boelens and Rob Kaptein for their useful remarks and criticism in our regular group seminars. Beyond that, I thank Alexandre Bonvin for the possibility that FANDAS became a part of WeNMR.

I continue with the actual group members, the people who surrounded me every day. The people who were there in the very beginning: Chris who introduced me to the world of biomolecular solid-state NMR, Dirk who shared the office and the fondness for French culture with me, Mikael K. for entertaining me with all his crazy jokes and Jacques for sharing his immense knowledge about NMR probeheads and hardware. Not to forget the people who were there (nearly) all the time. Marie - thanks for updating my French, our dinners together and being an amazing host in Marseille. Deepak - my physicist friend thanks for teaching me so much about NMR and our long chat sessions in the evenings and weekends. Markus, Elwin - thank you for being friends instead of simple colleagues. Abhi - if I mentioned everything I am grateful for, it would be a chapter of its own in this thesis. So I will keep it rather short here. Thank you for all your support in personal and professional matters and of course your excellent spicy Indian food which helped me to survive the Dutch cuisine. Barbara - the good soul of the department thank you for all your help and patience regarding administrative issues. Furthermore, I thank Marc for his support that FANDAS became a part of WeNMR and Johan for his help with all computer and spectrometer-related problems. Hans, Ram and Klaartje - thanks for making teaching

a great experience instead of a burden. In addition, I thank Hans for providing me a glimpse into solution-state NMR and Klaartje for the translation of the summary. And last but not least, I thank all the present and former members of the NMR spectroscopy group in Utrecht for keeping a nice and friendly atmosphere throughout the whole time.

I extend my thanks to people outside the UU circle.

Milena - my dearest friend thank you for all your wise words, inspiration and for constantly reminding me of the fact, that the world is not black and white but a colorful place.

Furthermore, I am grateful to my family and all my friends who stayed close to me and consoled me in rough times in spite of the distance. Warm thoughts go to Axel, Anna, Babs, Birgit, Nicki, Jakob, Verena, Sandra and Ulrich.

Finally, my four years would not have been the same without all the time I spent in Estonia. For all the fun, the relaxing, happy moments and the extraordinary people I was able to meet there, I thank the permanent members of the TSS team Aivar, Birgit, Dima, Gerli and Olga. Suur-suur aitäh 😊.

Financial support from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement n°211800 and NWO is gratefully acknowledged.

List of publications

The here present thesis is based on the following publications and manuscripts:

Gradmann S., Ader C., Heinrich I., Dittmann M., Cukkemane A., van Dijk M., Bonvin A. M., Engelhard M., Baldus M. *Rapid Prediction of multi-dimensional NMR data sets*. J. Biomol. NMR, **54**, 377-387 (2012).

Labohka A.A., **Gradmann S.**, Frey S., Hülsmann B. B., Urlaub H., Baldus M., Görlich D. *Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes*. EMBO J., **32**, 204-218 (2013).

Yang S., van Steenbergen M., Teunissen E., Novo L., **Gradmann S.**, Baldus M., van Nostrum C., Hennink W. *π - π Stacking increases the stability and loading capacity of thermosensitive polymeric micelles for chemotherapeutic drugs*. Biomacromolecules, **14**, 1826-1837 (2013).

Further publications:

Cukkemane A., Nand D., **Gradmann S.**, Weingarth M., Kaupp U.B., Baldus M. *Solid-state NMR [^{13}C , ^{15}N] resonance assignments of the nucleotide-binding domain of a bacterial cyclic nucleotide-gated channel*. Biomol. NMR Assign., **6**, 225-229 (2012).

Gradmann S., Medick P., Rössler E.A. *Glassy dynamics in nanoconfinement as revealed by ^{31}P NMR*. J. Phys. Chem. B, **113**, 8443-8445 (2009).

Curriculum vitae

Sabine Gradmann was born on March 1st, 1983 in Bayreuth, Germany. In 2002, she graduated from highschool with honors. From 2002 to 2008, she studied physics at the University of Bayreuth and Ecole Normale Supérieure de Lyon. During her studies, she received a stipend from the federal state of Bavaria. She started working in the field of solid-state NMR in her diploma thesis. There, she investigated the behavior of glass formers in nanoconfinement under the supervision of Prof. Ernst Rössler. Afterwards she continued as a research fellow in the same lab. In 2009, she started her PhD study at Utrecht University under the supervision of Prof. Marc Baldus. She received a Marie-Curie fellowship within the Initial Training Network (ITN) Structural Biology of Membrane Proteins (SBMP) for the first three years of her PhD study.