

Proton-Detected Solid-State NMR of the Cell-Free Synthesized α -Helical Transmembrane Protein NS4B from Hepatitis C Virus

Vlastimil Jirasko⁺,^[a] Nils-Alexander Lakomek⁺,^[a] Susanne Penzel,^[a] Marie-Laure Fogeron,^[b] Ralf Bartenschlager,^[c, d] Beat H. Meier,^{*[a]} and Anja Böckmann^{*[b]}

Proton-detected 100 kHz magic-angle-spinning (MAS) solid-state NMR is an emerging analysis method for proteins with only hundreds of microgram quantities, and thus allows structural investigation of eukaryotic membrane proteins. This is the case for the cell-free synthesized hepatitis C virus (HCV) nonstructural membrane protein 4B (NS4B). We demonstrate NS4B sample optimization using fast reconstitution schemes that enable lipid-environment screening directly by NMR. 2D

spectra and relaxation properties guide the choice of the best sample preparation to record 2D ^1H -detected ^1H , ^{15}N and 3D ^1H , ^{13}C , ^{15}N correlation experiments with linewidths and sensitivity suitable to initiate sequential assignments. Amino-acid-selectively labeled NS4B can be readily obtained using cell-free synthesis, opening the door to combinatorial labeling approaches which should enable structural studies.

Introduction

Membrane proteins in lipids can be studied by magic-angle-spinning (MAS) solid-state NMR^[1] which, different from solution NMR, does not face overall tumbling limitations. Thus the molecular weight of biomolecular complexes that can be addressed is solely limited by spectral resolution and sensitivity. Still, analysis faced a major obstacle: complex membrane proteins often need dedicated expression systems to functionally fold, but which may deliver low yields only and are thus often incompatible with the high sample quantities (> 10 mg) needed for classical ^{13}C -detected NMR approaches. This restricted NMR studies mostly to well-expressed model mem-

brane proteins, such as rhodopsin,^[1a,c,j,2] a truncated variant of influenza A M2 channel,^[1e,f] KcsA^[1d,k] or BmrA.^[3] Applications of the more sensitive ^1H -detected MAS NMR approach to membrane proteins emerged using extensively deuterated proteins and 10–20 kHz MAS frequency.^[4] Only recently, MAS frequencies around 100 kHz have allowed to study fully proton back-exchanged or even fully protonated proteins.^[5] The increase in MAS frequency resulted in a concomitant decrease in sample amount, to below the milligram,^[6] which presents a reduction in protein of about a factor 100 compared to ^{13}C detection, that is roughly compensated by the sensitivity gain inherent to ^1H -detection techniques.^[7] The first 60 to 100 kHz MAS schemes have been applied to well-established membrane protein systems,^[8] such as proteorhodopsin,^[9] outer membrane beta-barrels,^[1m,10,11] VDAC,^[11,12] a truncated variant of influenza A M2 channel,^[13] KcsA^[14] or BamA.^[15] Sample preparation of these proteins mainly followed previously established protocols using ^{13}C detection, and included formation of 2D crystals^[1m,12,13] or reconstitution into liposomes at low lipid-to-protein ratio (LPR, w/w) by dialysis.^[9,15] Such conditions are however not generally applicable to membrane proteins, and achieving optimal membrane reconstitution remains a critical step.

With sample amounts decreasing below the milligram, the use of eukaryotic protein expression systems becomes feasible for NMR sample preparation. In this context, the wheat-germ (WG) cell-free protein synthesis (CFPS) is a promising approach, since it provides one of the highest yields amongst eukaryotic CFPS systems, reaching routinely milligram amounts.^[16] WG-CFPS thus presents an efficient alternative to cellular eukaryotic expression systems, and has the advantage that various NMR isotope labeling schemes, including amino-acid selective labeling, can be easily implemented.^[16d,17] Importantly, also deuteration in combination with complete amide protonation


[a] Dr. V. Jirasko,⁺ Dr. N.-A. Lakomek,⁺ Dr. S. Penzel, Prof. Dr. B. H. Meier
ETH Zürich, Physical Chemistry
Vladimir-Prelog Weg 2, 8093 Zürich (Switzerland)
E-mail: beme@ethz.ch


[b] Dr. M.-L. Fogeron, Dr. A. Böckmann
Institut de Biologie et Chimie des Proteines
MMSB, Labex Ecofect, UMR 5086 CNRS, Université de Lyon
7 passage du Vercors, 69367 Lyon (France)
E-mail: a.boeckmann@ibcp.fr

[c] Prof. Dr. R. Bartenschlager
Department of Infectious Diseases
Molecular Virology, Heidelberg University
Im Neuenheimer Feld 345, 69120 Heidelberg (Germany)

[d] Prof. Dr. R. Bartenschlager
Division of Virus-Associated Carcinogenesis (Germany)
Cancer Research Center (DKFZ)
Im Neuenheimer Feld 242, 69120 Heidelberg (Germany)

[*] These authors contributed equally to this work.

 Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201900765>.

 © 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

can be achieved directly during synthesis,^[18] avoiding a denaturation and refolding step, which can compromise the native fold of a membrane protein.

We here investigated the nonstructural protein 4B (NS4B) of the hepatitis C virus (HCV). Around 70 million people are chronically infected with HCV and have a high risk to develop severe liver disease, including hepatocellular carcinoma. NS4B is a transmembrane protein that is essential for HCV genome replication and virion assembly.^[19,20] It has a sequence length of 261 amino acid (aa) residues (apparent molecular weight 27 kDa) and is an oligomeric α -helical transmembrane protein constituted of three subdomains.^[21] The central subdomain contains four predicted transmembrane segments.^[22] The N- and C-terminal subdomains each comprise two putative α -helices, presumably lying on the membrane surface. Current structural information is limited to the two isolated amphipathic helices located at the N terminus, AH1 (aa 4–32, PDB ID: 2LVG) and AH2 (aa 42–66, 2JXF) as well as to the C-terminal helix H2 (aa 229–253, 2KDR) which were all investigated by solution-state NMR on synthetic peptides representing the described helices.^[22,23] Overall, structural and topological information on full-length NS4B remains however sparse.^[24]

As NS4B is difficult to express in large quantities using conventional systems such as *Escherichia coli*, we established WG-CFPS for NS4B in a detergent-solubilized form.^[25] The protein can then be reconstituted into liposomes using Bio-Bead-enhanced dialysis.^[26] This is however a time-consuming step and simplifying this process is thus key to enable further sample optimization.

Here, we show how lipid reconstitution of NS4B can be optimized thanks to fast lipid-insertion schemes combined with direct screening using 2D ¹H-detected MAS-NMR spectra. We show that the achieved line-narrowing results from a decrease in inhomogeneous rather than homogeneous linewidth. Under the optimized conditions, 2D and 3D ¹H-detected correlation spectra can be recorded, both on fully and selectively labeled NS4B, initiating the crucial step of NMR sequential backbone assignments.

Results and Discussion

Cyclodextrin-mediated reconstitution yields NS4B proteo-liposomes

Previously, we have shown that NS4B can be expressed in a soluble form using WG-CFPS^[25a] and that further reconstitution into liposomes allowed to record quite well-resolved ¹³C-detected solid-state NMR spectra.^[25b] Resolution in spectra using the more sensitive ¹H detection remained however limited.^[25b] Efficient sample optimization was hampered by the lengthy dialysis step for lipid reconstitution and complete detergent removal.^[3a] In previous work, we adapted a faster method consisting of a combined detergent removal using cyclodextrin, as initially proposed by DeGrip et al.,^[27a] and proteoliposome separation on a sucrose gradient,^[27b,c] which speeded up the procedure by an order of magnitude. As in this approach parameters have to be fine-tuned to avoid protein loss,^[27b] we here

use a simplified approach (Figure 1 B) consisting of gradual addition of cyclodextrin to the detergent solubilized protein and lipids, followed by gradient centrifugation. This separation into two steps is indeed easier to handle and minimizes protein loss.

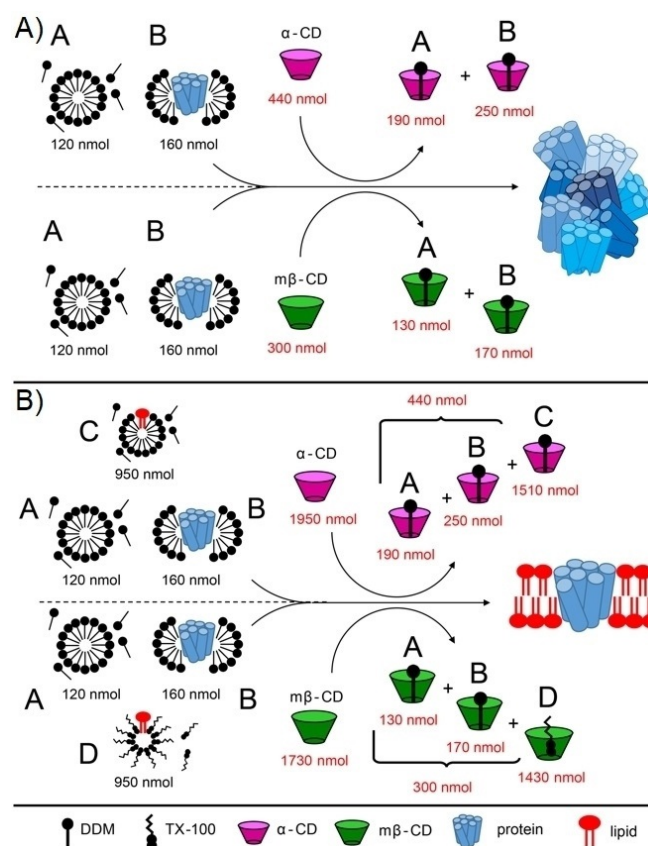


Figure 1. Cyclodextrin mediated lipid reconstitution of NS4B. Schematic representation of detergent removal by cyclodextrin from a 1 nmol solution of NS4B in 0.1% DDM buffer, A) in absence, and B) in presence of detergent-solubilized lipids. “A”, fraction of unbound DDM in a buffer; “B” fraction of DDM bound to protein; “C” and “D” DDM and Triton X-100 respectively used for lipid solubilization. Quantities are given in black and red for detergent and cyclodextrin, respectively (for raw data see Figure S1 and for calculations see Table S1). The detergent-to-lipid ratio in fractions “C” and “D” is 10, and the lipid-to-protein ratio is 2.

In order to assess the correct amount of cyclodextrin (CD) for protein reconstitution, we first determined the minimal amount of CD necessary to bind and remove detergent molecules from the *n*-dodecyl β -D-maltoside (DDM)-solubilized NS4B micelle complexes in the absence of lipids (Figure S1 a,c in the Supporting Information),^[27b,c] as outlined in Figure 1 A. Full precipitation of the protein was used as a read out for successful detergent removal. We tested two different cyclodextrins, α -cyclodextrin (α -CD) and methyl- β -cyclodextrin ($m\beta$ -CD), and found that a total of 110 nmol of α -CD fully precipitated 0.25 nmol of NS4B 0.1% DDM (Figure S1 a), corresponding to 440 nmol per 1 nmol of NS4B (Figure 1 A, upper panel). In case of $m\beta$ -CD, 75 nmol was used up per 0.25 nmol of NS4B (Figure S1 c), resulting in a total of 300 nmol per 1 nmol of NS4B (Figure 1 A lower panel).

In a second step, we estimated the amount of CD necessary to remove additional DDM or Triton X-100 detergents used for lipid solubilization. Experimental details of the precipitation assay in the presence of solubilized lipids or additional Triton X-100 are summarized in Figure S1 b and c, as well as in the Supporting results and Table S1. As a result (Figure 1B), 1950 nmol of α -CD is needed to reconstitute 1 nmol of NS4B into DDM-solubilized lipids, while 1730 nmol of $m\beta$ -CD is required for reconstitution into Triton X-100-solubilized lipids. From these experiments, the number of NS4B-bound DDM molecules was estimated as 160 (Supporting results and Table S1), in agreement with previously published data, where micelles with model membrane proteins contain around 150–300 DDM monomers.^[28] Further, we determine that α -CD binds at an approximate ratio of 1.6 to DDM, $m\beta$ -CD at 1.1 to DDM and $m\beta$ -CD at 1.5 to Triton X-100 (Supporting results and Table S1), which correlates with previous experiments showing that cyclodextrin molecules do not necessarily bind detergents in 1:1 ratio.^[27a,29]

We then prepared proteoliposomes using the above-described protocol and cyclodextrin amounts. The reconstituted protein was analyzed on a sucrose gradient, and a visible opaque band in the gradient was detected, which correlated with the expected protein bands of NS4B on an SDS PAGE gel (Figure S1 d, e). No further signals were detected by SDS PAGE, indicating that virtually all protein formed proteoliposomes. The protocol established here allowed to gain a factor of 10 in time, which also makes the reconstitution step more robust with respect to protein degradation.

After this biochemical assessment, we directly optimized the amount and type of cyclodextrin by NMR. To enhance readability of the spectra, we used a selectively Gly- and Tyr-labeled NS4B sample (dGY NS4B). We reconstituted NS4B into PC/Chol (phosphatidylcholine/cholesterol) at LPR 2 using different conditions, as detailed in Figure S2 and Table S2. As a result, the 2.5-fold excess of $m\beta$ -CD over the minimal amount determined, combined with PC/Chol solubilized in Triton X-100, yielded the spectra with best resolution and signal-to-noise ratio (SNR), and this condition was thus set as the sample preparation standard. Finally, 2D ^1H , ^{15}N correlation spectra (Figure S3) confirmed that NS4B lipid reconstitution using cyclodextrin indeed yielded very similar NMR spectra as the more time-consuming reconstitution using Bio-Bead-enhanced dialysis,^[25b] and could thus be used as starting point for further optimizations.

Proton linewidths depend on the lipid environment

Next, we compared spectra recorded at different LPRs to identify optimal conditions. For maximal protein NMR signal, the LPR, in principle, should be as low as possible. On the other hand, a sufficient amount of lipids is important to ensure that the protein is well folded.^[30] We therefore compared dGY NS4B reconstituted into PC/Chol lipids at various LPRs. The resulting 2D ^1H , ^{15}N correlation spectra (Figure 2) reveal that the spectral linewidth improves by about a factor of two from LPR 0.25 to LPR 4 (Figure S4, Table S3). Increase of LPR requires an increase

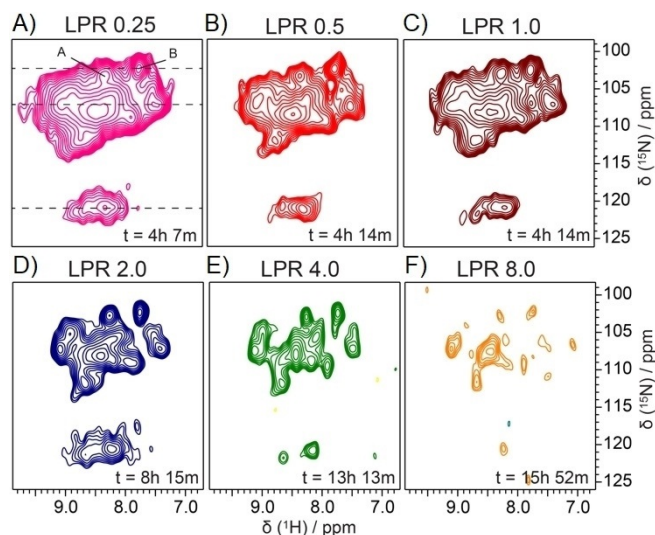


Figure 2. Spectral resolution of 2D ^1H , ^{15}N spectra as a function of LPR. Two-dimensional ^1H , ^{15}N correlation spectra of dGY NS4B reconstituted into PC/Chol liposomes at LPRs from 0.25 to 8. Two-dimensional ^1H , ^{15}N correlation spectra were acquired at 850 MHz, 80 kHz (A–D) or 90 kHz (E, F) MAS and 23 °C. Dashed lines represent the position of 1D traces shown in Figure S4. Lipid reconstitution was achieved by Bio-Bead-enhanced dialysis (A–C) or by using $m\beta$ -CD (D–F). Full width at half maximum (FWHM) and SNR are given in Table S3.

in measurement time, which becomes prohibitive at LPRs exceeding about 2. We therefore chose LPR 2 as the best compromise between a narrow spectral linewidth and high SNR (Figures 2 and S4, Table S3).

Not only LPR but also lipid membrane thickness, hydrophobic mismatch, the presence of cholesterol and the bilayer fluidity, among others, can influence protein folding and the dynamic properties of the protein in the membrane.^[31] We thus investigated the dependence of NMR spectral parameters on selected lipids. For this, [^2H , ^{15}N , ^{13}C] (dUL) NS4B as well as [^1H , ^{15}N , ^{13}C] (UL) NS4B was reconstituted at LPR 2 into different lipids as detailed in Figure 3.

To obtain a proxy for the spectral quality of NS4B in the different lipid environments, we determined the average ^1H FWHM and SNR for 9–10 isolated peaks (Table 1). A comparison (Figure 3 and Table 1) of linewidths showed that NS4B reconstituted into PC, both in the presence and absence of 30% cholesterol, as well as NS4B reconstituted into DMPC lipids, gave spectra with an average spectral linewidth between 100–120 Hz for isolated peaks. In contrast, NS4B reconstituted into DPPC lipids showed only poorly resolved spectra.

In solid-state NMR, the total linewidth (Δ^{total}) has a homogeneous (Δ^{homo}) and inhomogeneous (Δ^{inhomo}) contribution ($\Delta^{\text{total}} = \Delta^{\text{homo}} + \Delta^{\text{inhomo}}$). The homogeneous part represents coherent effects (Δ^{coherent}) caused by incomplete averaging of the dipolar interaction, and incoherent effects ($\Delta^{\text{incoherent}}$) driven by molecular dynamics, and thus $\Delta^{\text{homo}} = \Delta^{\text{coherent}} + \Delta^{\text{incoherent}}$. The inhomogeneous part reflects sample and field inhomogeneity.^[32] To determine the homogeneous contribution ($\Delta^{\text{homo}} = R_2'/\pi$) to the total linewidth we measured bulk ^1H R_2' relaxation-rate constants of amide groups using a Hahn-echo experiment.^[33] As

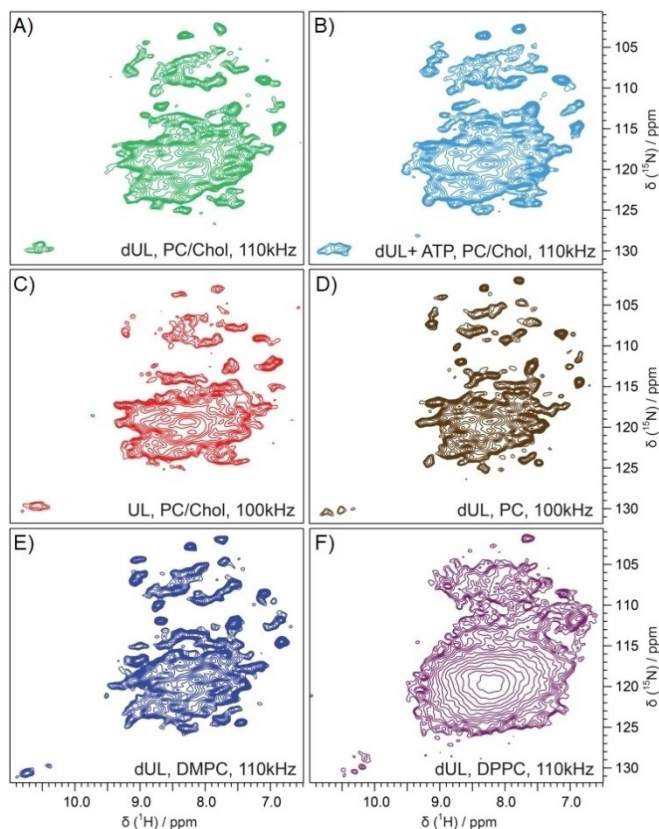


Figure 3. ^1H -detected 2D $^1\text{H},^{15}\text{N}$ spectra as a function of the composition of the lipid environment. A, B, D, E, F) dUL NS4B, or C) UL NS4B reconstituted in liposomes (LPR 2) with different lipid composition: A)–C) PC/Chol, D) PC, E) DMPC and F) DPPC. B) NS4B was reconstituted into lipids in the presence of ATP. Spectra were recorded at 850 MHz and MAS frequency of 110 kHz. All spectra were acquired at 23 °C.

Δ^{coherent} depends mainly on the geometry of the proton spin system,^[32] it should be similar in comparable secondary structure elements. In rigid, α -helical parts of deuterated (and HN back-exchanged) proteins, typical values of $\Delta^{\text{homogeneous}}$ are around 30 Hz at 100 kHz MAS in the absence of dynamics ($\Delta^{\text{incoherent}} \approx 0$).^[34] For dUL NS4B, we found values around 60–80 Hz (Table 1) suggesting that $\Delta^{\text{incoherent}}$ is about 30–50 Hz, which in-

dicates that the linewidth may be influenced by dynamics. This value derived for $\Delta^{\text{incoherent}}$ is similar to the measured ^1H $R_{1\rho}$ transverse relaxation rate constant (measured at a spinlock field of 13 kHz and tabulated in Table 1 as the rate constant divided by π and thus representing a linewidth). Δ^{total} is broadest for the protonated systems, but the differences are small, which might be due to the significant experimental error bars. We also measured ^{15}N $R_{1\rho}$ and R_2' rate constants (Table 1), which are, as expected, much lower than for protons. Interestingly, a comparison of $\Delta^{\text{homogeneous}}$ of DUL NS4B showed only small differences for the different lipid environments, including in the badly resolved spectra of NS4B in DPPC lipids (Table 1). Therefore, the difference in total linewidth between NS4B in DPPC (Figure 3F) and in the other lipids (Figure 3A, D, E) must be predominantly due to inhomogeneous line broadening and might be correlated to the different phase-transition temperatures of the various lipids (Table 1, see also discussion below).

To assess whether the observed linewidth is related to the lipid phase-transition temperature (T_m), we recorded spectra of NS4B in different lipid environments and at different temperatures (Figure S5). While for NS4B in PC liposomes, which show a T_m of -6°C ,^[35] the temperature dependence seems only weak (Figure S5 A,C), for DMPC, with a T_m of 24°C , the spectral resolution clearly decreases from 21 to -6°C (Figure S5 B, D). This suggests that spectral resolution is indeed influenced by the experimental temperature relative to T_m . This is similar to previous findings for the β -barrel membrane protein OmpX in lipid nanodiscs.^[10b] A comparison of $\Delta^{\text{homogeneous}}$ and also the ^1H $R_{1\rho}/\pi$ rate constants indicate only very little difference between the two temperatures, independent of the lipid environment (Table S4). Thus, in conclusion, the difference in spectral resolution cannot be explained by dynamics-related homogeneous line broadening alone (Table S4), but inhomogeneous line broadening below the lipid phase transition temperature has an important contribution as well.

We also recorded the spectra of NS4B prepared in the presence of ATP. NS4B is able to hydrolyze ATP,^[36] which might be essential for NS4B function in the HCV life cycle.^[37] We thus speculated that the addition of ATP during the reconstitution step could affect NS4B folding and improve sample quality.

Table 1. Homogeneous and inhomogeneous ^1H and ^{15}N linewidths of NS4B in different lipid environments.

Labeling	Lipids	T_m ^[a] [°C]	hNH ^[b] Figure 3	MAS ^[c] [kHz]	SNR ^[d]	^1H			^{15}N		
						Δ^{total} FWHM ^[e] [Hz]	$\Delta^{\text{homogeneous}}$ R_2'/π ^[f] [Hz]	$\Delta^{\text{inhomogeneous}}$ FWHM- R_2'/π ^[g] [Hz]	$R_{1\rho}/\pi$ ^[h] [Hz]	R_2'/π ^[i] [Hz]	$R_{1\rho}/\pi$ ^[j] [Hz]
dUL	PC/Chol		A	110	7 ± 2	120 ± 20	60 ± 10	60 ± 30	29 ± 1	14 ± 1	3 ± 1
dUL/ATP	PC/Chol		B	110	9 ± 3	120 ± 30	60 ± 10	60 ± 30	30 ± 1	15 ± 1	3 ± 1
UL	PC/Chol		C	110	4.5 ± 1.2	140 ± 30	150 ± 10	0 ± 30	174 ± 6	12 ± 1	3 ± 1
dUL	PC	-6	D	100	8 ± 2	100 ± 20	70 ± 10	30 ± 30	34 ± 2	18 ± 1	11 ± 1
dUL	DMPC	24	E	110	11 ± 3	110 ± 30	60 ± 10	50 ± 30	36 ± 1	15 ± 1	5 ± 1
dUL	DPPC	41	F	110	–	–	80 ± 10	–	26 ± 1	24 ± 1	3 ± 1

[a] " T_m ", lipid phase-transition temperature; [b] "hNH", panel in Figure 3 with the respective spectrum; [c] "MAS", spinning frequency; [d] "SNR", average SNR for 9 or 10 isolated peaks (for individual values see Table S5); [e] " Δ^{total} ", average ^1H FWHM of the same peaks; [f] " $\Delta^{\text{homogeneous}}$ ", bulk ^1H homogeneous linewidth as measured in a Hahn-echo t_2' experiment (Figure S7); [g] " $\Delta^{\text{inhomogeneous}}$ ", bulk ^1H inhomogeneous linewidth calculated from $\Delta^{\text{total}} - \Delta^{\text{homogeneous}}$; [h] " $R_{1\rho}/\pi$ ", proton rotating-frame relaxation time extracted from a bulk ^1H $R_{1\rho}$ experiment (Figure S7); [i] ^{15}N " R_2'/π " and [j] ^{15}N " $R_{1\rho}/\pi$ ", ^{15}N bulk relaxation-rate constants calculated by taking the inverse of T_2' and $T_{1\rho}$ times that were fitted from intensity decay curves (Figure S7).

Chemical shifts and linewidths in $^1\text{H},^{15}\text{N}$ correlation spectrum of NS4B-ATP sample were however comparable to the sample prepared in the absence of ATP (Figure 3B). At the same time, approximately 20% improvement of SNR as a direct or indirect effect of ATP was observed, which resulted in selection of this sample for 3D experiments (see below).

3D experiments on fully and selectively labeled NS4B for sequential backbone assignments

A set of ^1H -detected 3D experiments at 110 kHz MAS was acquired to evaluate the suitability of the dUL NS4B sample for sequential backbone assignment, a crucial step for further structural NMR analysis of the protein. Three spectra, hCANH, hCONH and hCAcoNH,^[6a,34,38] were recorded at 110 kHz MAS and an hCBcaNH spectrum^[34,39] was acquired at 60 kHz MAS frequency in a 1.3 mm rotor (Table S6).

Out of 248 expected resonances, prolines and flexible C-terminal tag not counted, approximately 190, 155 and 100 resonances were picked by the automatic peak-picking routine of the CCPN software package in the hCANH, hCONH and hCAcoNH spectra, respectively. Out of the 27 expected Gly residues, if nine residues in the flexible tag are neglected, 22 resonances in the Gly spectral region were visible in the hCANH spectrum. Although 3D spectra still show signal overlap, more than 60 peaks in the hCANH spectrum could be connected to their counterparts in the hCONH spectrum, as shown for a number of examples in Figure S6.

To reduce the significant signal overlap in central regions of the spectra, the recording of 4D spectra^[40] would be useful. For SNR reasons we instead turned to selectively labeled samples to decongest the spectra and obtain anchor points for the sequential backbone assignments. We prepared a deuterated

Gly, Val and Leu selectively labeled NS4B sample (dGVL NS4B). Three-dimensional hCANH and hCONH spectra were recorded and their ^{13}C projections are shown on the $^1\text{H},^{15}\text{N}$ dUL NS4B planes in Figure 4B and C, respectively. Although, in principle, all labeled residues in the sequence will contribute resonances in the hCANH spectrum, for the hCONH spectrum only pairs of either Gly, Val and Leu will contribute to a resonance signal, as both the $C'(i-1)$ and the $N(i)$ have to be isotopically labeled.

The automated peak picking selected 44 peaks in the hCANH spectrum, out of 74 expected resonances. Eleven resonances were picked in an hCONH spectrum out of expected 24 intraresidue correlations. Eight out of 11 correlations in the dGVL NS4B hCONH spectrum could be connected to residues in the hCANH dUL NS4B spectrum. Out of those, two Gly and one Leu could be assigned, namely L123, G125 and G238. One signal in the hCONH spectrum was assigned to a double Gly motif, which can be found once in the NS4B sequence and five times in the twin strep tag sequence. It is likely that this single GG pair belongs to NS4B rather than the tag, which is presumably flexible and therefore invisible in a CP-based experiment.

Combining information from the 3D hCANH, hCONH, hCAcoNH and hCBcaNH spectra^[34] of uniformly and selectively labeled samples, we were able to identify two amino-acid stretches, comprising residues from Val119 to Gly125 (VV¹²⁰SGLVG) and from Leu237 to Thr241 (LGSL²⁴⁰T). Corresponding strip plots are shown in Figure 5. $\Delta\delta_{\text{C}\alpha}-\Delta\delta_{\text{C}\beta}$ secondary chemical-shift differences of the amino-acid stretch VV¹²⁰SGLV gave positive values (Figure 5B), suggesting α -helical secondary structure, in agreement with previously proposed topological models^[22,36] (Figure 5C). On the other hand, within the amino-acid stretch LGSL²⁴⁰T, a positive $\Delta\delta_{\text{C}\alpha}-\Delta\delta_{\text{C}\beta}$ secondary chemical-shift difference for N-terminal leucine is observed, whereas Ser, Leu and Thr show values close to 0

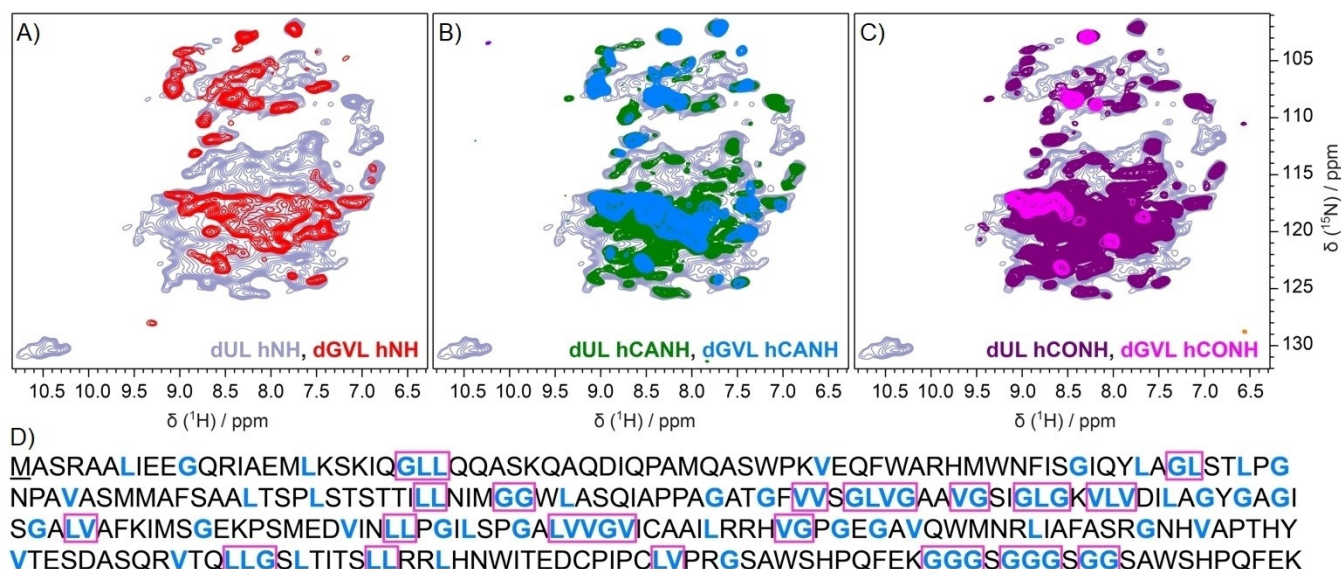


Figure 4. Selectively labeled dGVL NS4B. $^1\text{H},^{15}\text{N}$ 2D correlation spectra of dUL NS4B reconstituted at LPR 2 with ATP (in grey) overlaid with A) the dGVL NS4B $^1\text{H},^{15}\text{N}$ 2D correlation spectrum in red, B) the projection of the 3D hCANH spectra of dUL NS4B in green and dGVL NS4B in blue, and C) the projection of the 3D hCONH spectra of dUL NS4B in violet and dGVL NS4B in magenta. D) NS4B sequence with Gly, Val and Leu amino acids highlighted. Consecutive pairs that result in expected resonances in the 3D hCONH spectrum are shown in magenta boxes. The N-terminal methionine (attached for translation initiation) as well as the C-terminal thrombin cleavage site, GSA linker and tandem *Strep*-tag II are underlined.

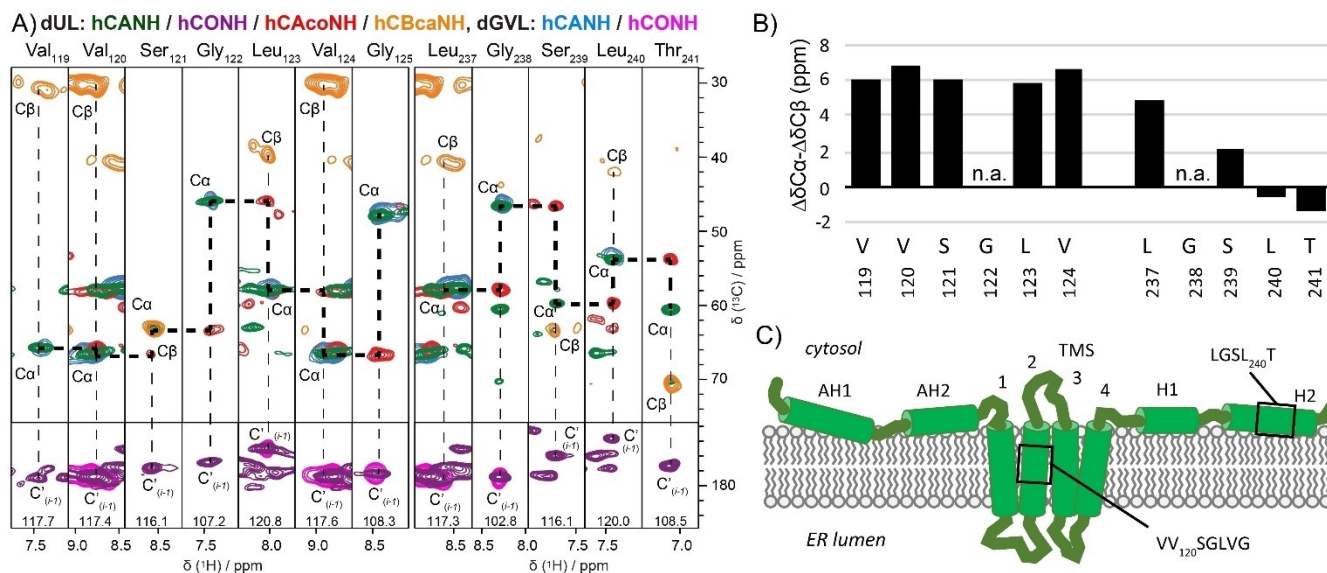


Figure 5. 3D correlation spectra of deuterated NS4B allow to establish sequential connectivities. A) Selected strip plots representing the assignment of amino acid stretches $VV_{120}SGLVG$ and $LGSL_{240}T$ using 3D hCANH (green), hCONH (violet) and hCAcoNH (red) spectra recorded on dUL NS4B, and hCANH (cyan) and hCONH (magenta) of dGVL NS4B in PC/Chol lipids LPR 2, at 110 kHz MAS. The dUL NS4B hCBcaNH spectrum (orange) was acquired at 60 kHz MAS. B) Secondary chemical shift differences, $\Delta\delta_{C\alpha}-\Delta\delta_{C\beta}$ of the two stretches from (A). C) Putative topology model of the NS4B protein adapted from Gouttenoire et al.^[22] in which NS4B is proposed to contain eight α -helices. N- and C terminus harbor four presumably amphipathic α -helices: AH1 (aa 4–32, PDB ID: 2LVG), AH2 (aa 42–66, 2JXF), H1 (predicted between aa 201–212) and H2 (aa 229–253, 2KDR). The central part (aa 70–190) contains four predicted transmembrane segments. The black boxes indicate the location of the two assigned amino acid stretches $VV_{120}SGLVG$ and $LGSL_{240}T$, respectively. Random-coil chemical shifts used for the calculation of secondary chemical shifts were taken from Zhang et al.^[41]

(Figure 5B), suggesting no defined secondary structure for that part. This is in conflict with a topological model based on a solution-state NMR analysis of an isolated peptide, which suggested full α -helical character for the entire region.^[22,23b] However, our data might support a more complex picture which was proposed for full-length NS4B, in which the C-terminus containing a putative walker B motif has been described to interact with the loop between *trans*-membrane helices 2 and 3, comprising a walker A motif.^[36]

Further spectral backbone assignment is currently challenged by a lack in spectral resolution and sensitivity. Indeed, the high spectral overlap of the mainly α -helical NS4B resonances, in combination with significant transverse relaxation and concomitant low signal to noise due to the membrane insertion at relatively high LPR, are problematic for less sensitive experiments, such as for example, the hCBcaNH experiment, and also experiments which allow for the identification of sequential $N(i)$ and $N(i-1)$ connectivities.^[42,43] Although such experiments are feasible in principle, their overall low transfer efficiency prevents data collection with present-day equipment. Currently, recording of 3D hCANH and hCONH spectra on different selectively labeled samples is therefore the most promising approach to obtain further sequential backbone assignments. Sophisticated combinatorial labeling schemes devised in the context of solution NMR studies^[44] can inspire further solid-state NMR approaches for sequential backbone assignment employing a combination of selectively labeled samples. Still, also here SNR is an issue, and higher magnetic field strength should boost the SNR and reduce signal overlap in the near future. Even faster magic angle spinning will increase

transfer efficiencies not only in CP as used here, but also in J coupling-based experiments. For instance, changing from 100 kHz to 200 kHz MAS frequency will elongate the transverse t_2' times by roughly a factor of 2^[32,45] and possibly overcompensate the SNR loss by the smaller sample in faster and therefore smaller rotors.

Conclusion

We show for the cell-free synthesized α -helical integral membrane protein NS4B of HCV in membranes that solid-state NMR spectra could be recorded in a reasonable amount of measurement time on a membrane protein sample reconstituted in lipids at LPR 2 in a 0.7 mm rotor. We screened for optimal sample preparation using rapid lipid reconstitution via cyclodextrin addition, and assessed the best lipid-to-protein ratio directly on 1H -detected solid-state NMR spectra. Relaxation measurements confirmed the expected narrower homogeneous linewidth of deuterated protein compared to protonated one, and revealed that inhomogeneous line broadening was substantial, and strongly dependent on the lipid chosen, which is likely related to the lipid phase transition temperature. The evaluation of different lipids showed that reasonably resolved spectra can be reproducibly recorded, and that most conditions yield similar spectra, with the notable exception of DPPC. Three-dimensional experiments were recorded and, in principle, provide the basis for sequential backbone assignments. Still, spectral overlap is substantial, and we showed that selectively labeled samples, straightforward to obtain by CFPS, can be used to identify anchor points for sequential assignments.

Further improvement in resolution is however compulsory to progress to complete backbone assignment, and with current hardware, the use of combinatorial labeling is thus the most promising approach.

Acknowledgements

This work was supported by an ERC Advanced Grant (B.H.M., grant no. 741863, *Faster*), by the Swiss National Science Foundation (B.H.M., grant no. 200020_159707), the French ANRS (France Recherche Nord & Sud, Sida- HIV et Hépatites), an autonomous agency at INSERM, France, the French Agence Nationale de Recherche (A.B., ANR-14-CE09-0024B) and the LABEX ECOFACT (A.B., ANR-11-LABX-0048) within the Université de Lyon program Investissements d'Avenir (A.B., ANR-11-IDEX-0007). We greatly acknowledge technical support by Riccardo Cadalbert, Andreas Hunkeler and Alexander Däpp (solid-state NMR group, ETH Zürich). Further, we would like to thank the entire Protein Solid-State NMR group, MMSB Lyon, for support in wheat-germ sorting.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: cell-free protein synthesis · lipid reconstitution · proton detection · solid-state NMR · transmembrane proteins

- [1] a) G. Gröbner, I. J. Burnett, C. Glaubitz, G. Choi, A. J. Mason, A. Watts, *Nature* **2000**, *405*, 810–813; b) D. Huster, X. L. Yao, M. Hong, *J. Am. Chem. Soc.* **2002**, *124*, 874–883; c) A. F. L. Creemers, S. Kihne, P. H. M. Bovee-Geurts, W. J. DeGrip, J. Lugtenburg, H. J. M. de Groot, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9101–9106; d) A. Lange, K. Giller, S. Hornig, M. F. Martin-Eauclaire, O. Pongs, S. Becker, M. Baldus, *Nature* **2006**, *440*, 959–962; e) S. D. Cady, K. Schmidt-Rohr, J. Wang, C. S. Soto, W. F. DeGrado, M. Hong, *Nature* **2010**, *463*, 689–692; f) M. Sharma, M. G. Yi, H. Dong, H. J. Qin, E. Peterson, D. D. Busath, H. X. Zhou, T. A. Cross, *Science* **2010**, *330*, 509–512; g) S. H. Park, B. B. Das, F. Casagrande, Y. Tian, H. J. Nothnagel, M. Chu, H. Kiefer, K. Maier, A. A. De Angelis, F. M. Marassi, S. J. Opella, *Nature* **2012**, *491*, 779–783; h) S. A. Shahid, B. Bardiaux, W. T. Franks, L. Krabben, M. Habeck, B. J. van Rossum, D. Linke, *Nat. Methods* **2012**, *9*, 1212–1217; i) S. L. Wang, V. Ladizhansky, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *82*, 1–26; j) S. L. Wang, R. A. Munro, L. C. Shi, I. Kawamura, T. Okitsu, A. Wada, S. Y. Kim, K. H. Jung, L. S. Brown, V. Ladizhansky, *Nat. Methods* **2013**, *10*, 1007–1012; k) B. J. Wylie, M. P. Bhate, A. E. McDermott, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 185–190; l) H. Kaur, A. Lakatos-Karoly, R. Vogel, A. Noll, R. Tampe, C. Glaubitz, *Nat. Commun.* **2016**, *7*, 13864; m) J. S. Retel, A. J. Nieuwkoop, M. Hiller, V. A. Higman, E. Barbet-Massin, J. Stanek, L. B. Andreas, W. T. Franks, B. J. van Rossum, K. R. Vinthkumar, L. Handel, G. G. de Palma, B. Bardiaux, G. Pintacuda, L. Emsley, W. Kuhlbrandt, H. Oschkinat, *Nat. Commun.* **2017**, *8*, 2073.
- [2] a) M. Etkorn, S. Martell, O. C. Andronesi, K. Seidel, M. Engelhard, M. Baldus, *Angew. Chem. Int. Ed.* **2007**, *46*, 459–462; *Angew. Chem.* **2007**, *119*, 463–466; b) J. Yang, L. Aslimovska, C. Glaubitz, *J. Am. Chem. Soc.* **2011**, *133*, 4874–4881.
- [3] a) B. Kunert, C. Gardiennet, D. Lacabanne, D. Calles-Garcia, P. Falson, J. M. Jault, B. H. Meier, F. Penin, A. Böckmann, *Front. Mol. Biosci.* **2014**, *1*, 5; b) D. Lacabanne, C. Orelle, L. Lecoq, B. Kunert, C. Chuilon, T. Wiegand, S. Ravaud, J.-M. Jault, B. H. Meier, A. Böckmann, *Commun. Biol.* **2019**, *2*, 149.
- [4] a) R. Linser, M. Dasari, M. Hiller, V. Higman, U. Fink, J. M. L. del Amo, S. Markovic, L. Handel, B. Kessler, P. Schmieder, D. Oesterhelt, H. Oschkinat, B. Reif, *Angew. Chem. Int. Ed.* **2011**, *50*, 4508–4512; *Angew. Chem.* **2011**, *123*, 4601–4605; b) M. E. Ward, L. Shi, E. Lake, S. Krishnamurthy, H. Hutchins, L. S. Brown, V. Ladizhansky, *J. Am. Chem. Soc.* **2011**, *133*, 17434–17443.
- [5] a) M. J. Knight, A. L. Webber, A. J. Pell, P. Guerry, E. Barbet-Massin, I. Bertini, I. C. Felli, L. Gonnelli, R. Pierattelli, L. Emsley, A. Lesage, T. Herrmann, G. Pintacuda, *Angew. Chem. Int. Ed.* **2011**, *50*, 11697–11701; *Angew. Chem.* **2011**, *123*, 11901–11905; b) M. Huber, S. Hiller, P. Schanda, M. Ernst, A. Böckmann, R. Verel, B. H. Meier, *ChemPhysChem* **2011**, *12*, 915–918.
- [6] a) V. Agarwal, S. Penzel, K. Szekely, R. Cadalbert, E. Testori, A. Oss, J. Past, A. Samoson, M. Ernst, A. Böckmann, B. H. Meier, *Angew. Chem. Int. Ed.* **2014**, *53*, 12253; *Angew. Chem.* **2014**, *126*, 12450; b) L. B. Andreas, K. Jaudzems, J. Stanek, D. Lalli, A. Bertarello, T. Le Marchand, D. C. D. Paepe, S. Kotelovica, I. Akopjana, B. Knott, S. Wegner, F. Engelke, A. Lesage, L. Emsley, K. Tars, T. Herrmann, G. Pintacuda, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9187–9192.
- [7] a) L. Lecoq, M. Schledorn, S. Wang, S. Smith-Penzel, A. A. Malär, M. Callon, M. Nassal, B. H. Meier, A. Böckmann, *Front. Mol. Biosci.* **2019**, *6*, 58; b) V. S. Mandala, M. Hong, *Curr. Opin. Struct. Biol.* **2019**, *29*, 30058.
- [8] T. Schubeis, T. Le Marchand, L. B. Andreas, G. Pintacuda, *J. Magn. Reson.* **2018**, *287*, 140–152.
- [9] D. Lalli, M. N. Idso, L. B. Andreas, S. Hussain, N. Baxter, S. Han, B. F. Chmelka, G. Pintacuda, *J. Am. Chem. Soc.* **2017**, *139*, 13006–13012.
- [10] a) O. Saurel, I. Iordanov, G. Nars, P. Demange, T. Le Marchand, L. B. Andreas, G. Pintacuda, A. Milon, *J. Am. Chem. Soc.* **2017**, *139*, 1590–1597; b) N. A. Lakomek, L. Frey, S. Bibow, A. Böckmann, R. Riek, B. H. Meier, *J. Phys. Chem. B* **2017**, *121*, 7671–7680.
- [11] E. E. Najbauer, K. T. Movellan, T. Schubeis, T. Schwarzer, K. Castiglione, K. Giller, G. Pintacuda, S. Becker, L. B. Andreas, *ChemPhysChem* **2019**, *20*, 302–310.
- [12] M. T. Eddy, Y. C. Su, R. Silvers, L. Andreas, L. Clark, G. Wagner, G. Pintacuda, L. Emsley, R. G. Griffin, *J. Biomol. NMR* **2015**, *61*, 299–310.
- [13] L. B. Andreas, M. Reese, M. T. Eddy, V. Gelev, Q. Z. Ni, E. A. Miller, L. Emsley, G. Pintacuda, J. J. Chou, R. G. Griffin, *J. Am. Chem. Soc.* **2015**, *137*, 14877–14886.
- [14] S. Jekhmene, J. Medeiros-Silva, J. Li, F. Kummerer, C. Muller-Hermes, M. Baldus, B. Roux, M. Weingarth, *Nat. Commun.* **2019**, *10*, 123.
- [15] C. Pinto, D. Mance, T. Sinnige, M. Daniels, M. Weingarth, M. Baldus, *Nat. Commun.* **2018**, *9*, 4135.
- [16] a) B. E. Roberts, B. M. Paterson, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 2330–2334; b) H. R. Pelham, R. J. Jackson, *Eur. J. Biochem.* **1976**, *67*, 247–256; c) K. Takai, T. Sawasaki, Y. Endo, *Nat. Protoc.* **2010**, *5*, 227–238; d) D. Lacabanne, M. L. Fogeron, T. Wiegand, R. Cadalbert, B. H. Meier, A. Böckmann, *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *110*, 20–33.
- [17] a) D. A. Vinarov, B. L. Lytle, F. C. Peterson, E. M. Tyler, B. F. Volkman, J. L. Markley, *Nat. Methods* **2004**, *1*, 149–153; b) M. Tonelli, K. K. Singarapu, S. Makino, S. C. Sahu, Y. Matsubara, Y. Endo, M. Kainosho, J. L. Markley, *J. Biomol. NMR* **2011**, *51*, 467–476.
- [18] S. Wang, M. L. Fogeron, M. Schledorn, M. Dujardin, S. Penzel, D. Burdette, J. M. Berke, M. Nassal, L. Lecoq, B. H. Meier, A. Böckmann, *Front. Mol. Biosci.* **2019**, *6*, 9187–9111.
- [19] V. Lohmann, S. Hoffmann, U. Herian, F. Penin, R. Bartenschlager, *J. Virol.* **2003**, *77*, 3007–3019.
- [20] a) M. Elazar, P. Liu, C. M. Rice, J. S. Glenn, *J. Virol.* **2004**, *78*, 11393–11400; b) H. Lindström, M. Lundin, S. Häggström, M. A. A. Persson, *Virus Res.* **2006**, *121*, 169–178; c) D. M. Jones, A. H. Patel, P. Targett-Adams, J. McLauchlan, *J. Virol.* **2009**, *83*, 2163–2177.
- [21] a) G. Y. Yu, K. J. Lee, L. Gao, M. M. Lai, *J. Virol.* **2006**, *80*, 6013–6023; b) J. Gouttenoire, P. Roingard, F. Penin, D. Moradpour, *J. Virol.* **2010**, *84*, 12529–12537; c) D. Paul, I. Romero-Brey, J. Gouttenoire, S. Stoitsova, J. Krijnse-Locker, D. Moradpour, R. Bartenschlager, *J. Virol.* **2011**, *85*, 6963–6976.
- [22] J. Gouttenoire, R. Montserret, D. Paul, R. Castillo, S. Meister, R. Bartenschlager, F. Penin, D. Moradpour, *PLoS Pathog.* **2014**, *10*, e1004501.
- [23] a) J. Gouttenoire, V. Castet, R. Montserret, N. Arora, V. Raussens, J. M. Ruyschaert, E. Diesis, H. E. Blum, F. Penin, D. Moradpour, *J. Virol.* **2009**,

- 83, 6257–6268; b) J. Gouttenoire, R. Montserret, A. Kennel, F. Penin, D. Moradpour, *J. Virol.* **2009**, *83*, 11378–11384.
- [24] J. Gouttenoire, F. Penin, D. Moradpour, *Rev. Med. Virol.* **2010**, *20*, 117–129.
- [25] a) M. L. Fogeron, A. Badillo, V. Jirasko, J. Gouttenoire, D. Paul, L. Lancien, D. Moradpour, R. Bartenschlager, B. H. Meier, F. Penin, A. Böckmann, *Protein Expression Purif.* **2015**, *105*, 39–46; b) M. L. Fogeron, V. Jirasko, S. Penzel, D. Paul, R. Montserret, C. Danis, D. Lacabanne, A. Badillo, J. Gouttenoire, D. Moradpour, R. Bartenschlager, F. Penin, B. H. Meier, A. Böckmann, *J. Biomol. NMR* **2016**, *65*, 87–98.
- [26] J. L. Rigaud, D. Levy, G. Mosser, O. Lambert, *Eur. Biophys. J.* **1998**, *27*, 305–319.
- [27] a) W. J. DeGrip, J. Vanoostrum, P. H. Bovee-Geurts, *Biochem. J.* **1998**, *330*, 667–674; b) D. Lacabanne, A. Lends, C. Danis, B. Kunert, M. L. Fogeron, V. Jirasko, C. Chuilon, L. Lecoq, C. Orelle, V. Chaptal, P. Falson, J. M. Jault, B. H. Meier, A. Böckmann, *J. Biomol. NMR* **2017**, *69*, 81–91; c) T. Althoff, K. M. Davies, S. Schulze, F. Joos, W. Kuhlbrandt, *Angew. Chem. Int. Ed.* **2012**, *51*, 8343–8347; *Angew. Chem.* **2012**, *124*, 8468–8472.
- [28] V. Chaptal, F. Delolme, A. Kilburg, S. Magnard, C. Montigny, M. Picard, C. Prier, L. Monticelli, O. Bornert, M. Agez, S. Ravaud, C. Orelle, R. Wagner, A. Jawhari, I. Broutin, E. Pebay-Peyroula, J. M. Jault, H. R. Kaback, M. le Maire, P. Falson, *Sci. Rep.* **2017**, *7*, 41751.
- [29] G. A. Signorell, T. C. Kaufmann, W. Kukulski, A. Engel, H. W. Remigy, *J. Struct. Biol.* **2007**, *157*, 321–328.
- [30] F. Hagn, M. Etzkorn, T. Raschle, G. Wagner, *J. Am. Chem. Soc.* **2013**, *135*, 1919–1925.
- [31] D. Huster, *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2014**, *1841*, 1146–1160.
- [32] S. Penzel, A. Oss, M. L. Org, A. Samoson, A. Böckmann, M. Ernst, B. H. Meier, *J. Biomol. NMR* **2019**, *73*, 19–29.
- [33] E. L. Hahn, *Phys. Rev.* **1950**, *80*, 580–594.
- [34] S. Penzel, A. A. Smith, V. Agarwal, A. Hunkeler, M. L. Org, A. Samoson, A. Böckmann, M. Ernst, B. H. Meier, *J. Biomol. NMR* **2015**, *63*, 165–186.
- [35] R. Koynova, M. Caffrey, *Biochim. Biophys. Acta* **1998**, *1376*, 91–145.
- [36] A. A. Thompson, A. Zou, J. Yan, R. Duggal, W. Hao, D. Molina, C. N. Cronin, P. A. Wells, *Biochemistry* **2009**, *48*, 906–916.
- [37] S. Einav, M. Elazar, T. Danieli, J. S. Glenn, *J. Virol.* **2004**, *78*, 11288–11295.
- [38] E. Barbet-Massin, A. J. Pell, J. S. Retel, L. B. Andreas, K. Jaudzems, W. T. Franks, A. J. Nieuwkoop, M. Hiller, V. Higman, P. Guerry, A. Bertarello, M. J. Knight, M. Felletti, T. Le Marchand, S. Kotelovica, I. Akopjana, K. Tars, M. Stoppini, V. Bellotti, M. Bolognesi, S. Ricagno, J. J. Chou, R. G. Griffin, H. Oschkinat, A. Lesage, L. Emsley, T. Herrmann, G. Pintacuda, *J. Am. Chem. Soc.* **2014**, *136*, 12489–12497.
- [39] a) S. Xiang, J. Biernat, E. Mandelkow, S. Becker, R. Linser, *Chem. Commun.* **2016**, *52*, 4002–4005; b) P. Fricke, V. Chevelkov, M. Zinke, K. Giller, S. Becker, A. Lange, *Nat. Protoc.* **2017**, *12*, 764–782.
- [40] S. K. Vasa, H. Singh, P. Rovo, R. Linser, *J. Phys. Chem. Lett.* **2018**, *9*, 1307–1311.
- [41] H. Zhang, S. Neal, D. S. Wishart, *J. Biomol. NMR* **2003**, *25*, 173–195.
- [42] S. Q. Xiang, K. Grohe, P. Rovo, S. K. Vasa, K. Giller, S. Becker, R. Linser, *J. Biomol. NMR* **2015**, *62*, 303–311.
- [43] L. B. Andreas, J. Stanek, T. Le Marchand, A. Bertarello, D. C. D. Paepe, D. Lalli, M. Krejčíková, C. Doyen, C. Oster, B. Knott, S. Wegner, F. Engelke, I. C. Felli, R. Pierattelli, N. E. Dixon, L. Emsley, T. Herrmann, G. Pintacuda, *J. Biomol. NMR* **2015**, *62*, 253–261.
- [44] a) B. Hoffmann, F. Lohr, A. Laguerre, F. Bernhard, V. Dotsch, *Prog. Nucl. Magn. Reson. Spectrosc.* **2018**, *105*, 1–22; b) F. Lohr, J. Gebel, E. Henrich, C. Hein, V. Dotsch, *J. Magn. Reson.* **2019**, *302*, 50–63.
- [45] A. Böckmann, M. Ernst, B. H. Meier, *J. Magn. Reson.* **2015**, *253*, 71–79.

Manuscript received: December 17, 2019

Accepted manuscript online: December 18, 2019

Version of record online: February 20, 2020