

Supporting Information

Fast Magic-Angle-Spinning NMR Reveals the Evasive Hepatitis B Virus Capsid C-Terminal Domain

M. Callon, A. A. Malär, L. Lecoq, M. Dujardin, M.-L. Fogeron, S. Wang, M. Schledorn, T. Bauer, M. Nassal, A. Böckmann, B. H. Meier**

Material and Methods

Sample preparation

[¹³C, ¹⁵N] uniformly labelled Cp183 capsids were prepared in *E. coli* as described in ^[1] and sedimented into Bruker 1.3 mm and 3.2 mm MAS rotors via ultracentrifugation (200,000 *g*, 14 h, 4 °C) using a home-made filling device $[2]$.

[¹H, ¹³C, ¹⁵N]-arginine-labelled HBV capsids were prepared in *E. coli* using an unlabelled and fully deuterated ($[^{2}H, ^{12}C, ^{14}N]$) medium as follows: for each preparation, 1 liter of M9 medium was prepared in 100 % D_2O and supplemented with 1 g of ${}^{2}H-{}^{12}C$ -glucose-D7, 1 g of ${}^{14}NH_4Cl$ and 0.5 g of ${}^{13}C^{-15}N$ -Arginine (Eurisotop). Plasmids of pET-28a2-HBc149opt (for Cp149), pRSF_T7-HBc183opt (for Cp183) and pRSF_Tet-SRPK1dNS1_T7-HBc183opt (for P7-Cp183) were transformed into *E. coli* BL21* CodonPlus (DE3) cells and grown at 37 °C for 7 hours in 50 ml LB medium containing chloramphenicol (34 µg/ml) and carbenicillin (50 µg/ml) for pET-28a2 vector or kanamycin (50 µg/ml) for pRST vectors. 20 ml from the LB preculture were centrifuged (1500 *g*, 10 min, 20 °C) and bacteria were gently resuspended in 200 ml of previously-described M9 medium with appropriate antibiotics and incubated overnight at 25 °C. Then, 800 ml of fresh M9 medium were added to the 200 ml and cells were grown at 37 °C for 8 h until OD_{600nm} reached \sim 2. Protein expression was induced with 1 mM isopropyl ß-D-1-thiogalactopyranoside (IPTG) overnight at 16 °C for P7-Cp183, 20 °C for Cp183 and 25 °C for Cp149. Cells were harvested by centrifugation (6000 *g*, 10 min, 4 °C). The three proteins were purified as described in [1]. For Cp149 and P7-Cp183, an additional gel filtration step was performed using a HiPrepTM 16/60 Sephacryl® S-200 HR column (120 ml) to get rid of bound Triton-X100 used during purification $[1]$. For Cp183, capsids were either used directly after purification (Cp183 filled with *E. coli* RNA), or disassembled and reassembled either with *in vitro*-transcribed pgRNA or with commercial torula yeast RNA (Sigma) as described in $^{[1]}$. Final samples were buffer exchanged in D₂O buffer (50 mM TRIS-D₂O, pH 7.5) using 3 cycles and 10-times dilution steps of concentration using Amicon Ultra-centrifugal filter units (Merck, 50 kDa cutoff). In total, five samples were prepared with specificallylabelled arginine residues ($[1H, 13C, 15N]$) with all other residues deuterated but unlabelled ($[2H, 16]$ ¹²C, ¹⁴N]): Cp149, P7-Cp183, Cp183 reassembled with pgRNA and yeast RNA, and Cp183 filled with *E. coli* RNA (summarized in Figure S3). Capsids were sedimented and filled into Bruker 0.7 mm MAS rotors via ultracentrifugation (200,000 *g*, 14 h, 4 °C) using a home-made filling device [2].

NMR spectroscopy and data processing

The ¹³C-detected solid-state NMR experiments recorded at low temperature (LT) were carried out on the [¹³C, ¹⁵N] uniformly labelled Cp183 capsids on a Bruker AVANCE III wide-bore (WB) spectrometer operating at a ¹H resonance frequency of 600 MHz. 2D NCX experiments were measured at 40 kHz MAS frequency in a 1.3 mm LT MAS probe (Bruker) in combination with a LT-MAS Cooling System (Bruker) for bearing, drive and sample cooling air. The temperature was calibrated with a sample of KBr powder ^[3]. NCX spectra were recorded at a VTU cooling temperature of 207 K and 100 K, respectively corresponding to 243 K (-30 °C) and 105 K (- 168 °C) inside the rotor. The equivalent NCX spectrum at 4 °C was recorded on an 800 MHz Bruker AVANCE III WB spectrometer in a 3.2 mm probe (Bruker) at a 17.5 kHz MAS. VTU cooling temperature was set to -18 $^{\circ}$ C, corresponding to 4 $^{\circ}$ C in the rotor according to the chemical shift of the supernatant water signal $[2]$. In these latter conditions, 2D HC-INEPT spectra were recorded on [¹³C, ¹⁵N] uniformly labelled Cp149 and Cp183 capsids. A summary of acquisition parameters is given in the supplementary Table S1.

The ¹H-detected solid-state NMR experiments were recorded using a Bruker AVANCE III wide-bore (WB) spectrometer operating at a ¹H resonance frequency of 850 MHz on $[$ ¹H, ¹³C, 15 N]-arginine labelled Cp149, P7-Cp183, Cp183 and yCp183 in D₂O buffer, and internallyreferenced to DSS. All experiments on arginine-labelled samples were measured at 100 kHz MAS in a 0.7 mm triple-resonance probehead from Bruker Biospin. The sample temperature of around 21°C was determined from the relationship $T(^{\circ}C) = 455 - 90^{\circ}\delta_{H2O}$ described in [4], where δ_{H2O} denotes the chemical shift of the supernatant water signal $[2]$. All used pulse sequences are summarized in Figure S4.

EXSY experiments (pulse sequence shown in Figure S4a) were measured with τ_{mix} = 10, 20, 50, 75 and 100 ms. Resonance-specific ¹H transverse relaxation times T_2 ['] (¹H) and rotatingframe relaxation times $T_{1p}({}^{1}H)$ have been determined from modified versions of the EXSY experiments containing an additional Hahn-echo (Figure S4b) and a 13 kHz spin-lock block (Figure S4c) respectively. T_2 '(¹H) and T_{1p} ¹H) have been measured with eight 2D ¹H-¹H experiments with a τ_{mix} of 10 ms and varying echo time τ_{echo} ranging from 5 μ s to 10 ms or variable spin-lock length $\tau_{spin-lock}$ ranging from 5 μ s to 2 ms, respectively.

Heteronuclear 2D hCH experiments have been recorded with CP-based and (non-refocused) INEPT transfer steps (Figure S4d and f respectively). CP lengths were set to 250 and 300 µs $(\tau_{CP,a}$ and $\tau_{CP,b}$ contact times respectively). In the INEPT based experiments, the INEPT delays τ_1 and τ_2 were both set to 1 ms. In the INEPT-EXSY experiment (Figure S4g), τ_{mix} was set to 50 ms. A 102 ms pre-saturation block with 50 Hz irradiation strength on the water frequency was used in all 2D ¹H-¹H correlation spectra. The MISSISIPPI scheme ^[5] was used for solvent suppression in the CP-based hCH experiments, while no pre-saturation was used for the INEPT-based hCH experiments.

Further acquisition parameters are given in the Supplementary Tables 2 and 3.

Spectroscopic data have been processed using Topspin 3.5pl6 (Bruker Biospin) with zero filling to the double amount of data point and a shifted sine-bell apodization function in direct and indirect dimensions with $SSB = 2.5$. Acquisition in the direct dimension was cut to 12.9 ms. Pre-existing arginine resonance assignments were transferred from earlier work $[6,7]$ (BMRB accession number 27845), using CcpNmr Analysis 2.4.2 [8–10] Peak positions and peak intensities were then exported to MATLAB (MATLAB R2019a, 9.6.0.) for the relaxation analysis.

Relaxation analysis

Experimental intensities of diagonal and off-diagonal peaks in the EXSY experiments have been fitted as function of the relaxation delay using a mono-exponential decay model with two free parameters (functional form $A \cdot \exp(-\tau/T)$).

Error bars on relaxation times have been obtained using bootstrapping methods with 200 iterations and are shown as 2σ , where σ is the experimental error. Corresponding homogeneous linewidths and relaxation-rate constants have been obtained from the experimental relaxation times according to the formulas $\Delta^{\text{homo}}(1\text{H}) = \frac{R_2'(1\text{H})}{\pi} = 1/(\pi \cdot T_2'(1\text{H})$ and $\Delta_{1\rho}(1\text{H}) = \frac{R_{1\rho}(1\text{H})}{\pi} = 1/(\pi \cdot R_{1\rho}(1\text{H})$. Error bars thereon have been obtained via Gaussian error propagation.

The proton $T_{1\rho}$ relaxation-rate constant coming from the modulation of the ${}^{1}H-{}^{13}C$ dipole interaction can be approximated by:

$$
R_{1\rho}^{CH} = \left(\frac{\delta^D}{4}\right)^2 \left(\frac{1}{3}J(\omega_1 - 2\omega_r) + \frac{2}{3}J(\omega_1 - \omega_r) + \frac{2}{3}J(\omega_1 + \omega_r) + \frac{1}{3}J(\omega_1 + 2\omega_r)\right)
$$
(1)

assuming an on-resonance spin lock. $[11,12]$. We neglect, in rough approximation, the proton homonuclear interaction^[13]. δ^D denotes the anisotropy of the ¹H-¹³C dipolar-coupling tensor and ω_1 the amplitude of the spin-lock field which was set to $2\pi \cdot 13$ kHz in all our experiments. For the C_a-H_a dipole we assume $\delta^D = 2 \pi 46.6$ kHz. The spectral density function for a motion with a single correlation time τ and an amplitude *S* is defined as

$$
J(\omega) = \frac{2}{5}(1 - S^2) \frac{\tau}{1 + (\omega \tau)^2}.
$$
 (2)

 R_2^{incoh} is given, in the Redfield limit, by:^[11]

$$
R_2^{\text{NH}} = \left(\frac{\delta^{\text{D}}}{4}\right)^2 \left(\frac{4}{3} J(\omega_{\text{r}}) + \frac{2}{3} J(2\omega_{\text{r}}) + \frac{1}{2} J(\omega_{I} - \omega_{s}) + 3J(\omega_{s}) + \frac{3}{2} J(\omega_{I}) + 3J(\omega_{I} + \omega_{s})\right) \tag{3}
$$

Supplementary figures

Supplementary Figure 1. Proton-detected CP and INEPT spectra of Cp149 and Cp183 capsids. a) 2D hNH spectra recorded at 100 kHz MAS on deuterated and back-exchanged ²H-¹³C-¹⁵N Cp149 (orange, taken from $^{[6]}$) and ²H-¹³C-¹⁵N Cp183 (green, taken from (Wang et al., 2019)) **b)** ¹H-¹³C-INEPT spectra of ¹³C-¹⁵N-Cp149 (orange) and ¹³C-¹⁵N-Cp183 with *E. coli* RNA (green) recorded at 17.5 kHz. Additional resonances were assigned to RNA, which are labeled since they were produced in a 13 C- 15 N medium, as the capsid. In both CP and INEPT based experiments, no additional signals in the chemical shift region typical of Arginine chemical shifts were detected (upper part in panel a) and red circles in panel b)).

Supplementary Figure 2. Low temperature experiments on E.coli RNA-containing Cp183. 13C-detected NCO spectra recorded at 277 K (4 °C, green), 243 K (-30 °C, blue) and 105 K (-168 °C, red) on ¹³C-¹⁵N Cp183 capsids. **a)** 1D traces through the backbone (δ (¹⁵N) = 119 ppm) and the side chains (δ ¹⁵N) = 72 and 84 ppm) regions. NCO signal intensities have been scaled to each other on the N-CO backbone correlation at δ (¹⁵N) = 119 ppm (left panel in a)), allowing to follow changes specifically in the arginine side chains region (middle and right panels). The gain in peak volume is a factor of 2.4 and 3.2 for the Nn-Cz correlation, and a factor of 3.5 and 3.3 for the N ε -C ζ correlation at 243 K and 105 K, respectively. The volume was calculated using the parabolic fit method in CcpNmr. **b)** Overlay of the corresponding 2D NCO spectra. For experimental parameters, see Table S1.

Supplementary Figure 3. Overview of the different arginine residue labelling schemes. **a)** Cp149 and P7-Cp183 in absence of RNA, **b)** Cp183 reassembled with pgRNA or with yeast RNA, and c) Cp183 containing RNA from *E. coli*.

Supplementary Figure 4. Overview of all used pulse sequences. Color coding is based on type of transfer where green is Spin Diffusion/NOE, blue is cross-polarization (CP) and violet is INEPT. **a-c)** EXSY based pulse sequences used for the 1 H-1 H spectroscopy and the corresponding relaxation data shown in Figure 2-3. **d-f)** have been used for the heteronuclear correlation experiments shown in Figure 4, while **g)** is an INEPT-EXSY hybrid experiment, containing an hCH INEPT out-and back step, followed by a spin diffusion/NOE mixing step and its results are shown in Figure 5.

Supplementary Figure 5. ¹H-¹H EXSY spectra. EXSY spectra using 50 ms mixing time of **a)** Cp149, **b)** P7-Cp183 and **c)** Cp183-pgRNA recorded at 100 kHz MAS and 850 MHz proton frequency.

Supplementary Figure 6. Mixing time dependence in EXSY spectra. EXSY spectra of **a)** Cp149, b) P7-Cp183, c) Cp183-pgRNA in D₂O measured at increasing ¹H-¹H homonuclear mixing times of 20, 50 and 75 ms and **d)** corresponding extracted traces at $\delta_1(^1H) = 1.67$ ppm with zooms at the H $_{\beta}$ /H $_{\alpha\gamma}$ region. Note that at longer mixing times (> 20 ms), we start observing

an $H_{\beta/\gamma}$ – $H_{\alpha, NTD}$, cross peak whose peak maximum is however upshifted compared to the random coil position of the P7-Cp183 cross peak maximum, indicating that it is indeed originating from transfers within the arginine residues of the NTD and that the corresponding cross-peak in the P7-Cp183 spectrum is related to the arginine residues in the flexible CTD.

Supplementary Figure 7. Plot of the diagonal and crosspeak intensities normalized with their maximal intensity function of the mixing time for Cp149 (orange), P7-Cp183 (cyan) and Cp183-pgRNA (steel blue) in D_2O .

Supplementary Figure 8. Relaxation decay curves. Shown are the measure data points for the relaxation measurement results shown in Figure 2f and h for all five capsid samples.

Supplementary Figure 9. 2D-hCH CP-based spectra of **a)** Cp149 (orange), **b)** P7-Cp183 (cyan) and **c)** Cp183pgRNA (steel blue) in D₂O buffer recorded with τ_{CP} = 250 and 300 µs at 100 kHz MAS and 850 MHz proton frequency. The NTD assignment is transferred from [6].

Supplementary Figure 10. Simulation of INEPT and refocused INEPT efficiencies. Calculations are based on the average relaxation data shown in Figure 2 for Cp149 and P7- Cp183. One clearly sees in the CH case the need to use short INEPT delays in the order of 1 ms (and the non-refocused version of the experiment).

Supplementary Figure 11. 2D-hCH INEPT-based spectra of **a)** Cp149 (orange) and **b)** Cp183-pgRNA (steel blue) in D₂O buffer recorded with $\tau_H = \tau_N = 1$ ms at 100 kHz MAS and 850 MHz proton frequency with 80 scans. The NTD assignment is transferred from [6]. The random coil chemical shift positions are shown with red circles $[14]$.

Supplementary Figure 12. INEPT peaks coincide with EXSY positions. Resonances of P7-Cp183 appear at random coil positions. Overlay of extracted trace at $\delta_1(^1H) = 1.67$ ppm of EXSY spectrum shown in Figure 2b (cyan) with extracted traces at $\delta_1(^{13}C) = 27.20$ ppm (orange), $\delta_1(^{13}C)$ = 30.22 ppm (purple), and $\delta_1(^{13}C)$ = 43.27 ppm (blue) of 2D hCH INEPT-based spectrum shown in Figure 4.

Supplementary Figure 13. Cp183 and yCp183 INEPT spectra. 2D-hCH INEPT-based spectra of **a)** Cp183 (green) and **b)** yCp183 (purple) samples in D₂O, recorded with $\tau_1 = \tau_2 = 1$ ms at 100 and 97 kHz MAS frequency respectively and 850 MHz proton frequency. The assignment is transferred from $[6,15]$.

Supplementary Figure 14. R133 spin system. Overlay of 2D hHC rINEPT-TOSBY experiment (green) with a 2D hCH rINEPT-based experiment (orange) of Cp149 sample in D₂O for identification of the R133 residue. The connections are highlighted with black dashed lines.

Supplementary Figure 15. Position of R133 in the capsid structure. a) zoom of the prolinerich loop (128T-136N) and helix 5 (yellow) of the NTD domain of the HBV capsid (PDB 1QGT). The two residues involved in a hydrogen bond are labeled and the hydrogen bond is show as yellow dotted line **b)**.

Supplementary tables

Supplementary Table 1. Experimental table for solid-state NMR acquisition and pulse program parameters used for the INEPT and NCX experiments.

Supplementary Table 2. Experimental table for solid-state NMR acquisition and pulse program parameters used for the EXSY experiments.

Supplementary Table 3. Experimental table for solid-state NMR acquisition and pulse program parameters used for the heteronuclear correlation experiments.

References

- [1] L. Lecoq, S. Wang, M. Dujardin, P. Zimmermann, L. Schuster, M.-L. Fogeron, M. Briday, M. Schledorn, T. Wiegand, L. Cole, R. Montserret, S. Bressanelli, B. H. Meier, M. Nassal, A. Böckmann, *Proceedings of the National Academy of Sciences* **2021**, *118*, e2022464118.
- [2] A. Böckmann, C. Gardiennet, R. Verel, A. Hunkeler, A. Loquet, G. Pintacuda, L. Emsley, B. H. Meier, A. Lesage, *Journal of Biomolecular NMR* **2009**, *45*, 319–327.
- [3] K. R. Thurber, R. Tycko, *Journal of Magnetic Resonance* **2009**, *196*, 84–87.
- [4] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *The Journal of organic chemistry* **1997**, *62*, 7512–7515.
- [5] D. H. Zhou, C. M. Rienstra, *Journal of Magnetic Resonance* **2008**, *192*, 167–172.
- [6] L. Lecoq, M. Schledorn, S. Wang, S. Smith-Penzel, A. A. Malär, M. Callon, M. Nassal, B. H. Meier, A. Böckmann, *Frontiers in Molecular Biosciences* **2019**, *6*, 58.
- [7] L. Lecoq, S. Wang, T. Wiegand, S. Bressanelli, M. Nassal, B. H. Meier, A. Böckmann, *Biomolecular NMR Assignments* **2018**, *12*, 205–214.
- [8] R. Fogh, J. Ionides, E. Ulrich, W. Boucher, W. Vranken, J. P. Linge, M. Habeck, W. Rieping, T. N. Bhat, J. Westbrook, K. Henrick, G. Gilliland, H. Berman, J. Thornton, M. Nilges, J. Markley, E. Laue, *Nature Structural Biology* **2002**, *9*, 416–418.
- [9] T. J. Stevens, R. H. Fogh, W. Boucher, V. A. Higman, F. Eisenmenger, B. Bardiaux, B. J. Van Rossum, H. Oschkinat, E. D. Laue, *Journal of Biomolecular NMR* **2011**, *51*, 437– 447.
- [10] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins-Structure Function And Bioinformatics* **2005**, *59*, 687–696.
- [11] U. Haeberlen, J. S. Waugh, *Physical Review* **1969**, *185*, 420–429.
- [12] S. Penzel, A. Oss, M. L. Org, A. Samoson, A. Böckmann, M. Ernst, B. H. Meier, *Journal of Biomolecular NMR* **2019**, *73*, 19–29.
- [13] P. Rovó, R. Linser, *The journal of physical chemistry B* **2017**, *121*, 6117–6130.
- [14] D. S. Wishart, B. D. Sykes, *Journal of Biomolecular NMR* **1994**, *4*, 171–180.
- [15] L. Lecoq, *Biomolecular NMR assignments* **2018**, *12*, 205-214.