Biophysical Journal, Volume *120*

Supplemental information

Relationships between RNA topology and nucleocapsid structure in a model icosahedral virus

Laurent Marichal, Laetitia Gargowitsch, Rafael Leite Rubim, Christina Sizun, Kalouna Kra, Stéphane Bressanelli, Yinan Dong, Sanaz Panahandeh, Roya Zandi, and Guillaume Tresset

Manuscript submitted to **Biophysical***Journal*
Article

Supplementary Material: Relationships between RNA Topology and Nucleocapsid Structure in a Model Icosahedral Virus

Laurent Marichal[1], Laetitia Gargowitsch[1], Rafael Leite Rubim[1], Christina Sizun[2], Kalouna Kra[1,3], Stéphane Bressanelli[3], Yinan Dong[4], Sanaz Panahandeh[4], Roya Zandi[4], and Guillaume Tresset[1,*]

[1]Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France
[2]Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, 91198 Gif-sur-Yvette, France
[3]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France
[4]Department of Physics and Astronomy, University of California, Riverside, California 92521, United States
[*]Correspondence: guillaume.tresset@universite-paris-saclay.fr

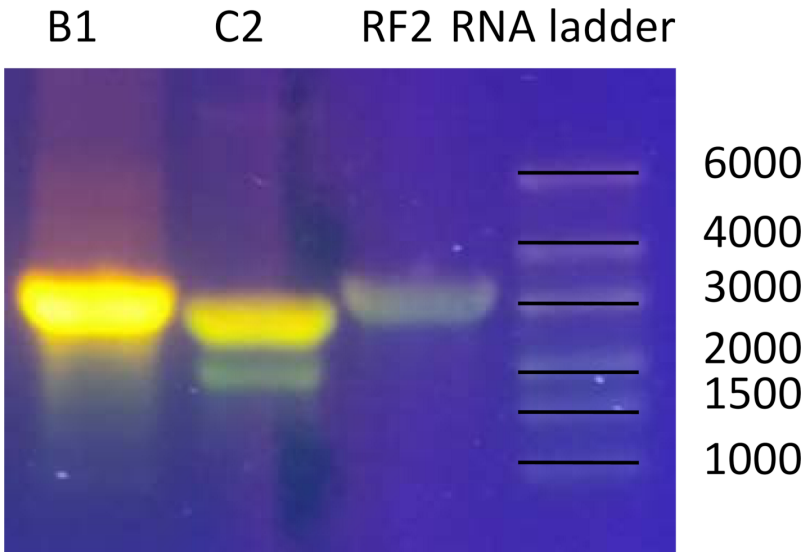

Figure S1: Agarose gel electrophoresis of the *in vitro* transcripted RNAs. The sizes of the B1, C2, and RF2 transcripts were compared with a RiboRuler High Range RNA Ladder (Thermo Fisher Scientific). Large bands located around the 3000 bases mark can be seen. This confirms the expected sizes of 3234, 2767, and 2687 bases, respectively. For C2, another band located at around 2000 bases can be seen but was deemed negligible compared to the intensity of the main band.

BINDING MEASUREMENTS BY NUCLEAR MAGNETIC RESONANCE

Concentration in subunits, i.e., dimeric CPs, was set to 37 μM and C2 concentration was varied from 0 to 110 nM. Resulting nucleoprotein complexes were dispersed in 80 mM NaCl, 20 mM Tris-HCl pH 7.5. Measurements were done in 3 mm tubes filled with 200 μl of sample and 10 μL ^2H_2O to lock the spectrometer frequency. ^{1}H nuclear magnetic resonance (NMR) spectra were collected at a temperature of 298 K on a 950-MHz NMR spectrometer (Bruker) equiped with a triple resonance TCI cryoprobe, using water suppression by excitation sculpting (zgesgp) and 2048 scans. Under these conditions, the signal of C2 was not detected. Due to the molecular weight of CP subunits (40 kDa), the observed signal is predominantly due to highly dynamic and disordered regions of CP. Addition of 110 nM C2 to 37 μM dimeric CP resulted in nearly complete loss of protein signal due to the formation of a very high molecular weight complex. C2 concentrations from 27.5 to 82.5 nM resulted in partial signal intensity loss. 500 mM NaCl restored signal intensity, suggesting complex disruption at high ionic strength.

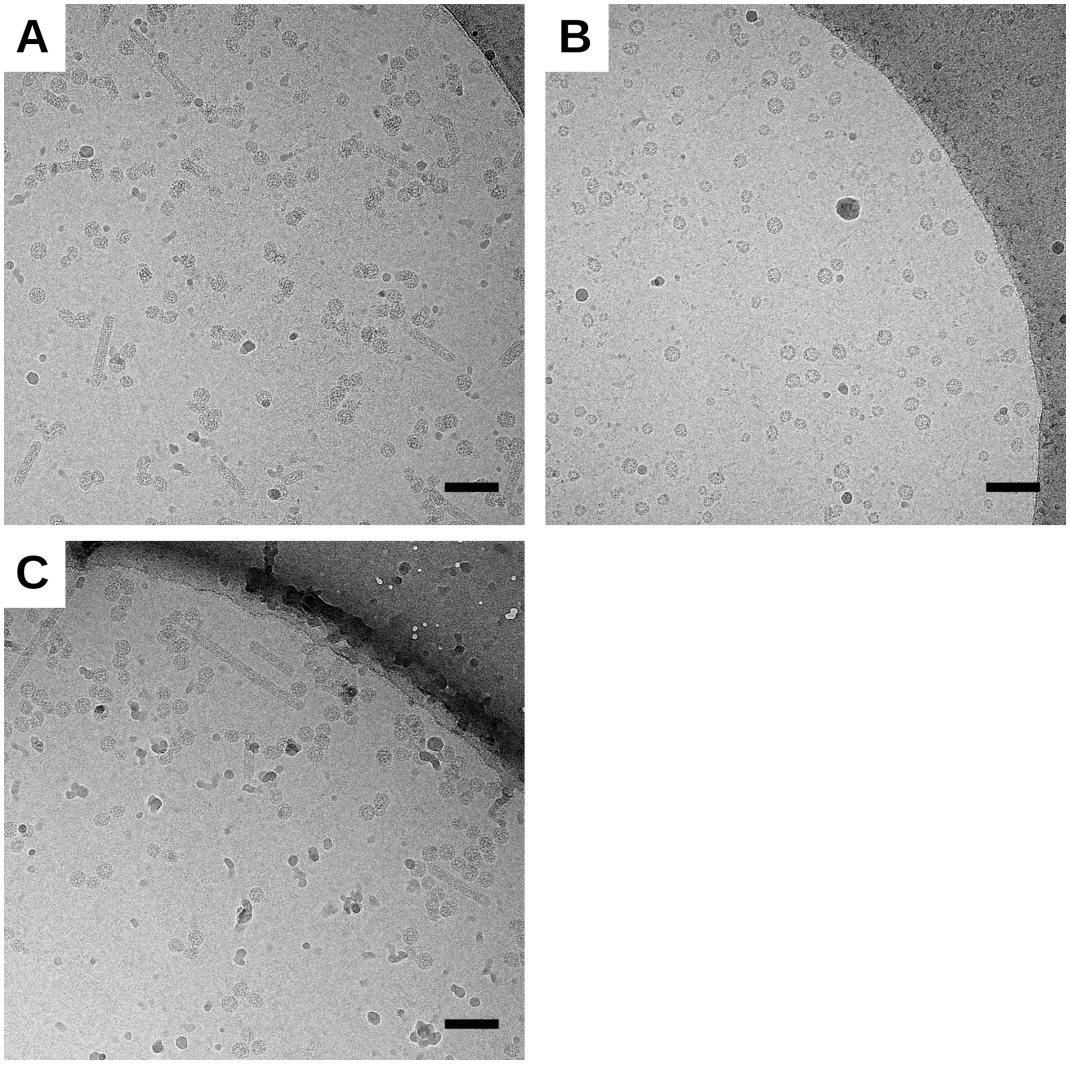

Figure S2: CryoTEM images of nucleocapsids with B1 (A), C2 (B) and RF2 (C). CP concentrations were 4 g.L^{-1}, 4 g.L^{-1} and 3.6 g.L^{-1}, while CP-to-RNA mass ratios were 6, 6 and 5.5, respectively. Scale bars are 100 nm.

The protein ^{1}H signal was integrated in the methyl (-0.413 to 0.985 ppm) as well as in the amide/aromatic proton (7.101 to 10.076 ppm) regions. The signal was normalized by that of free CP. Two analytical models were applied to predict the normalized integrated signal over . In both cases, subunits were assumed to exchange between their bound and free states. The fraction of free subunits f is given by

$$f = 1 - \frac{1}{2}\left[1 + \frac{K_{\mathrm{D}}}{c_{\mathrm{S}}} + r - \sqrt{\left(1 + \frac{K_{\mathrm{D}}}{c_{\mathrm{S}}} + r\right)^{2} - 4r}\right] \tag{1}$$

where K_{D} stands for the dissociation constant, c_{S} the subunit concentration and r the molar ratio between C2 binding sites and subunits, assuming that each RNA has 90 independent binding sites. In the slow exchange model, the normalized integrated signal varies directly as the fraction of free subunits, i.e., $I_{\mathrm{slow}} = f$.

In the fast exchange model, the mean relaxation rate is expressed by a linear combination of the relaxation rates corresponding to free ($R_{2\mathrm{free}}$) and bound ($R_{2\mathrm{bound}}$) subunits, and the normalized integrated signal becomes

$$I_{\mathrm{fast}} = \exp\left[-\tau(f.R_{2\mathrm{free}} + (1-f).R_{2\mathrm{bound}})\right] \tag{2}$$

where τ is a relaxation delay introduced by the zgesgp pulse sequence (~9 ms).

^{1}H R_2 relaxation rates were estimated taking into account dipole-dipole relaxation by another ^{1}H within a 2.6 Å distance (1). We assumed that the bound form adopted the same dynamic behavior as a capsid. Using a mean size of 250 Å and the Stokes-Einstein relation, the correlation time for overall rotational diffusion of a capsid was estimated to be 2000 ns, resulting in an $R_{2\text{bound}}$ value of the order of 1000 s^{-1}. For the free form, a rotational correlation time in the ns range, as observed for disordered proteins (2), gave $R_{2\text{free}}$ values in the 1-10 s^{-1} range.

Figure S3 depicts integrated ^{1}H NMR signals along with the signals predicted by the two models. It turned out that the fast exchange model (exchange rate between bound and free states > $R_{2\text{bound}}$) reproduced more closely the data with realistic values for the parameters (K_D = 1 μM), which was not the case with the slow exchange model (exchange rate < $R_{2\text{free}}$). However, we cannot exclude a third intermediate case, where the exchange rate would be comprised between 2 s^{-1} and 1000 s^{-1}.

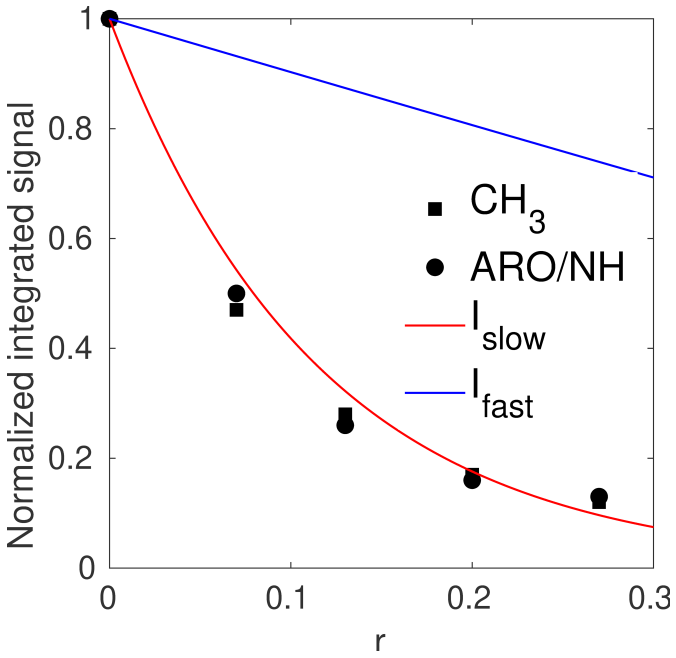

Figure S3: NMR normalized integrated signal as a function of the ratio of binding sites to subunits r. Measurements are represented by black squares and discs for methyl and amide/aromatic proton regions, respectively. These data were modeled with the slow (blue line) and fast (red line) exchange models. K_D = 1 μM, c_S = 37 μM, $R_{2\text{bound}}$ = 1000 s^{-1}, $R_{2\text{free}}$ = 2 s^{-1} and τ = 9 ms.

Table S1: CP and RNA concentrations of the samples used in the SAXS experiments. After dialysis in a moderate ionic strength buffer (50 mM NaCl, 50 mM Tris-HCl pH 7.5), concentrations are deduced from spectrophotometric measurements and calculated with the method developed by Porterfield and Zlotnick (3). The estimation of the polydispersity of the core radius ($\Delta R/R$) is based on a vesicle model.

	c_{CP} (g.L^{-1})	c_{RNA} (g.L^{-1})	CP-to-RNA mass ratio	$\Delta R/R$ (%)
Native virions	2.2	0.59	3.7	20.4±0.04
CP-B1	0.4	0.09	4.5	28.3±0.1
	1.2	0.29	4.1	26.1±0.07
	2.0	0.50	4.0	25.8±0.04
CP-C2	0.40	0.11	3.5	31.6±0.2
	0.94	0.30	3.1	26.0±0.05
	2.4	0.68	3.5	25.8±0.04
CP-RF2	0.40	0.05	7.5	N/A
	0.85	0.16	5.3	N/A
	3.6	0.43	8.3	N/A

Table S2: Size analysis of the cryoTEM 2D classification for nucleocapsids. The classes are numbered from left to right on Fig. 4 of the main text and nanotubes have been discarded for clarity. Particles are fitted with an ellipse of major radius a and minor radius b, a/b is therefore the aspect ratio. A native virion is expected to have $a = 16$ nm and $a/b = 1.0$.

	Class	a (nm)	a/b
CP-B1	1	16	1.02
	3	15	1.06
	4	15	1.15
	5	14	1.11
CP-C2	1	9.9	1.00
	2	16	1.01
	3	14	1.12
	4	16	1.25
	5	13	1.12
CP-RF2	1	15	1.01
	2	14	1.16
	3	13	1.09

REFERENCES

1. Sudmeier, J. L., S. E. Anderson, and J. S. Frye, 1990. Calculation of nuclear spin relaxation times. *Concepts Magn. Res.* 2:197–212.

2. Abyzov, A., N. Salvi, R. Schneider, D. Maurin, R. W. Ruigrok, M. R. Jensen, and M. Blackledge, 2016. Identification of dynamic modes in an intrinsically disordered protein using temperature-dependent NMR relaxation. *J. Am. Chem. Soc.* 138:6240–6251.

3. Porterfield, J. Z., and A. Zlotnick, 2010. A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance. *Virology* 407:281–288.