Supporting Information

Self-Assembly of Measles Virus Nucleocapsid-like Particles: Kinetics and RNA Sequence-Dependence

Sigrid Milles*, Malene Ringkjøbing Jensen*, Guillaume Communie†, Damien Maurin, Guy Schoehn, Rob W. H. Ruigrok,* and Martin Blackledge*

anie_201602619_sm_misceleaneous_information.pdf
Cloning of N⁰P fusion constructs.
The N-terminal fragment of the MeV phosphoprotein (residues 1-40, 1-50, or 1-100, Edmonston strain; UniProt Q83623) in fusion with residues 1-405, 1-525 or 30-405 of MeV nucleoprotein (Edmonston strain; UniProt Q89933) were cloned between NdeI and Xho1 cleavage sites of a pET41c (+) vector following a two-step PCR reaction. The genes coding for P and N were first amplified separately with an overlap between their 3’ and 5’ ends respectively. If desired, a TEV cleavage site (ENLYFQG) was included in the amplification primers to be able to separate P and N proteins after purification. The two DNA fragments were then annealed, and the fusion construct was submitted to another round of PCR amplification yielding constructs P₅₀N₄₀₅, P₅₀N₅₂₅, P₄₀₆₃₀₄₀₅ and P₁₀₀N₄₀₅. The final sequence of P₅₀N₅₂₅ was:

MAEEQARHVKNGLECIRALKAEPIGSLAIEEAMAAWSEISDNPGQERATCENLYFQGMATLLRSLALFKRNKDPPITSGSGGAIARGIKIIIIIIVIPGDSSITTRSRLDLRLVRLIGNPDVSGPKLTGALIGIIILFVESPGQLIQRTDDPVDSIRLLEVQVQDSQSQGLTFASRGTNMEDEADQYFSSHDDPISSDQSRFGWENKESDIEVQDPEFGFMILGTAQIWVLLAKAVTAPDIADELRRWIKYTTQQRRVVGFRLERKWLVDVVRNRIAEDSLRRFVMVALILDKRTPGNKPIAEMCIDDITYVEAGLASFILTISFRGQETMYPALGLHEFAGELSTLESLMNLQQMGETAPYMVILENSIQNKFSAGSYPLLWSYAMGVVELENSMGLNFGRSYFPDPAYFRGLQEMVRRASGKVSSTELGITAEDARLVSEIAMHTTEDKISRAGVPQAQVSLFHLHDQSENELPGKEDRRVQKSREGAREYSRTPGSRASDARAAHLPHTGLPTLDIDTAQESSQDPQDSRRSADALLRLQAMAGEESEQGSDTDTPIVYNDRLLDVHHHHHHHHH

Residues that stem from the TEV cleavage site or purification tag are underlined. Sequences of the other constructs were analogous. Only P₁₀₀N₄₀₅ did not contain a TEV cleavage site. All constructs were verified by DNA sequencing.

Expression and purification of N⁰P constructs.

Escherichia coli Rosetta™ (λDE3)/pRARE strain (Novagen) was used for the protein production. Cultures were grown at 37°C in lysogeny broth (LB) medium until the optical density (OD) at 600 nm reached 0.6. The temperature was then lowered to 20°C and expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested after 12-14 hours. For expression of labelled protein, 4 litres of culture were initially grown in LB medium and the cells were subsequently transferred into 1 litre of M9 medium at an OD₆₀₀ of 0.6. The cells were then grown for an additional 1 hour at 20°C before induction.

Cells were lysed by sonication in 20 mM Tris pH 8, 150 mM NaCl, 1 tablet Roche complete EDTA-free protease inhibitors, 1 spatula tip of lysozyme. Cell debris was harvested by centrifugation and the supernatant was loaded on Ni beads (His-select, Sigma Aldrich™). The flow through was discarded and beads were washed with 20 mM Tris pH 8, 150 mM NaCl, 8 mM imidazole. The protein was
eluted from the beads in 20 mM Tris pH 8, 150 mM NaCl, 400 mM imidazol. TEV cleavage was set up over night and at the same time the protein was dialyzed against 20 mM Tris pH 8, 150 mM NaCl, 5 mM beta-mercaptoethanol (BME). The protein was then purified by gel filtration (Superdex 200 column, GE Healthcare™) equilibrated with 50 mM phosphate pH 7, 150 mM NaCl, 5 mM BME.

**Electron microscopy**

The dependence of the assembly of NCLPs on RNA sequence was investigated by mixing $P_{50N_{405}}$ with RNA of different sequence composition: ACCAGA corresponding to six nucleotides from the 5’ end of the measles virus genome (5’-RNA$_6$), UUUUUU (polyU-RNA$_6$), AAAAAA (polyA-RNA$_6$) as well as a 60 nucleotide RNA from the 5’ end of the genome (5’-RNA$_{60}$) with sequence: ACCAGACAAAGCUGGAAUAGAAACUUCCGUAAUUUUCAAGUUUCUUUAUAUAUUG CAA

All RNAs were purchased from Integrated DNA Technologies, Leuven, Belgium and all were terminated with OH at both 3’ and 5’. The protein was mixed with the RNAs reaching final concentrations of 31 µM of $P_{50N_{405}}$ and 50 µM of RNA and incubated at room temperature for 24 hours. Four microliters of sample at approximately 0.1 mg/ml were adsorbed onto the clean face of a carbon film on a mica sheet (carbon/mica interface) and negatively stained with 2% (w/v) sodium silicotungstate (pH 7.5). Images were acquired with a FEI Tecnai12 LaB6 electron microscope working at 120 kV and with nominal magnifications of 35000 and 49000X, using a Gatan OriusTM SC1000 CCD camera.

The assembly of NC particles was studied in real time by electron microscopy by mixing $P_{50N_{405}}$ with 5’-RNA$_6$ reaching final concentrations of 22 µM of $P_{50N_{405}}$ and 50 µM of 5’-RNA$_6$. The sample was visualized at the following time points after mixing: 4 min, 9 min, 25 min and 67 min. In addition the $P_{50N_{405}}$ sample was visualized without RNA as a control showing no formation of NCLPs (data not shown).

The influence of the presence of $N_{TAIL}$ on the assembly of NCLPs was investigated by mixing $P_{50N_{525}}$ with either 5’-RNA$_6$ or 5’-RNA$_{60}$ reaching final concentrations of 20 µM of $P_{50N_{525}}$ and 50 µM of RNA.

**NMR spectroscopy**

The spectral assignment of $P_{50N_{405}}$, $P_{50N_{525}}$ and $P_{50N_{405}}$ in the presence of 5’-RNA$_{60}$ was obtained using $^{13}$C, $^{15}$N-labeled samples employing a set of BEST-TROSY triple resonance experiments$^{[2]}$ correlating $C_\alpha$, $C_\beta$ and $C^\prime$ resonances. The NMR experiments were acquired at 25°C in a buffer consisting of 50 mM phosphate, 150 mM NaCl, 5 mM BME at pH 7.0 on a Bruker spectrometer equipped with a cryo-probe operating at a $^1$H frequency of 700 MHz.
The assembly of NCLPs was followed by real-time NMR using a series of SOFAST $^1$H-$^{15}$N HMQC experiments using a sample of $^{15}$N-labeled N$^0$P complex (P$_{50N405}$ or P$_{50N525}$). The formation of NCLPs was initiated by addition of 5'RNA$_6$, 5'RNA$_60$, polyA-RNA$_6$, polyU-RNA$_6$ or 5'-DNA$_6$. A series of SOFAST HMQC$^{[3]}$ experiments were recorded with 100 complex points in the indirect dimension, a 200 ms recycling delay and four transients providing a time resolution of approximately four minutes. The spectra were recorded on a Bruker spectrometer operating at a $^1$H frequency of 950 MHz at 25°C in a buffer consisting of 50 mM phosphate, 150 mM NaCl, 5 mM BME at pH 7.0. In addition, two additional SOFAST HMQC experiments were recorded in the absence of RNA and at the end of the time course (after 24 hours) with 256 complex points in the indirect dimension and 16 transients (Fig. 3b, S5).

The experiments allow us to follow the evolution of the NMR spectra as P$_{50}$ is released from the N$^0$P complex and the NCLPs are formed. A time trace of the formation of the NCLPs were obtained by extracting the NMR signal intensities from each of the SOFAST HMQC spectra and plotted as function of time after addition of the RNA.

**NC assembly by fluorescence anisotropy**

P$_{50N405}$ or P$_{50N525}$ was diluted to the desired concentration into 50 mM Na-phosphate pH 7, 150 mM NaCl, 5 mM βME directly into the fluorescence cuvette. RNA$_{10}$-FAM (10 nuleotide RNA derived from the 5’ end of the MeV genome, with a fluorescein derivative attached to its 3’ end; ACCAGACAAA-FAM) was added to N$^0$P immediately prior to kinetics acquisition, mixed quickly through pipetting and fluorescence kinetics were recorded at an emission wavelength of 520 nm upon excitation with 470 nm light. Parallel and perpendicular polarization directions were recorded alternatingly and used pairwise to calculate the fluorescence anisotropy $r$

$$r = \frac{I_\parallel - GI_\perp}{I_\parallel + 2GI_\perp}$$

with the G-factor correcting for detection differences between parallel ($I_\parallel$) and perpendicular ($I_\perp$) polarized fluorescence light. G was determined on a daily basis according to standard protocols, and remained stable between days. Fluorescence spectra were recorded at the end of the assembly kinetics with an excitation wavelength of 460 nm and an emission range of 470-650 nm.

Assembly kinetics were analysed following exponential kinetics derived from a simple binding model assuming pseudo first order conditions:

$$r(t) = \sum_{i=0}^{n} \{A_i (1 - e^{-k_i t})\} + r_{t=0}$$

n was set to 2 depending on the requirements of the respective sample. A global fit was employed enforcing linearity between $k_i$ and N$^0$P concentration. The global fit for varying P$_{50N405}$ concentrations was achieved by considering the time between RNA addition and the start of the measurement, fixing
r_{t=0} to 0.05 and considering two independent rates (n=1). The apparent rate k_i in the fitting equation was replaced by k_i*·c_i, with k_i* the ith assembly rate constant and c_i the concentration of N^6P. c_i was fixed for the individual curves and the amplitudes A_i were fitted independently for the different N^9P concentrations. Due to the more complex assembly kinetics, the fit of P_{50}N_{525} NC assembly was simplified by keeping r_{t=0} as a fitting parameter per kinetic time trace. Only the two smallest protein concentrations (7.5 µM and 3.8 µM) required a third (fast) kinetic component in this case.
**Figure S1.** The three constructs of N and their nomenclature.
Figure S2. NMR-based identification of N_{405}-binding-region on the phosphoprotein P. (a) $^1$H-$^{15}$N HSQC of P_{1-100} (red) and P_{100}N_{405} (blue) superimposed. The inset shows the region of the tryptophan side chains (Trp36 of P_{1-100} is visible). (b) Intensity ratio of P_{100}N_{405} and P_{1-100} as obtained from the spectra shown in (a). The intensity ratio was normalised to a maximum of 1. Note that this P_{100}N_{405} construct does not contain a TEV cleavage site between P_{1-100} and N_{405}. 
Figure S3. Characterization of the P$_{50}$N$_{405}$ complex. (a) SDS-PAGE gel (Coomassie blue staining) showing the purity of P$_{50}$N$_{405}$ after cleavage by the TEV protease, Ni-affinity and size exclusion chromatography. P$_{50}$ being very small and weakly stained, only N is visible on the gel. Left lane shows molecular weight markers and numbers indicate the size of the respective bands (kDa). Note that resonances of P$_{50}$ residues are, however, clearly visible in our NMR spectra (see for example Figure S4). (b) Gel filtration profile of P$_{50}$N$_{405}$ showing the UV absorption profiles at 260 and 280 nm demonstrating that the complex does not contain RNA. (c) Size exclusion chromatography combined with detection by multi-angle laser light scattering (MALLS) and refractometry of P$_{50}$N$_{405}$ showing a mass of 53.7 kDa (expected mass for a heterodimeric complex: 52.6 kDa).
Figure S4.

Comparison of the $^1$H-$^{15}$N HSQC spectra of $P_{50}N_{405}$ and $P_{40}N_{30-405}$. Assignments of the resonances of $P_{50}N_{405}$ are shown. Note that $P_{50}$ contains six extra residues after cleavage from $N$ by TEV ($^{51}$ENLYFQ$^{56}$, subscript ‘T’ in the Figure) and that $N$ contains two extra residues ($^{406}$VE$^{407}$) before the non-cleavable his-tag used for purification. The main differences between the two spectra are, as expected, located at residues 44-56 of $P$ and residues 26-29 of $N$. 
Figure S5.

Comparison of $^1$H-$^{15}$N correlation spectra of $P_{50}N_{405}$ with and without RNA. Spectra are shown of $P_{50}N_{405}$ (red), $P_{50}$ (green) and $P_{50}N_{405}$ after incubation with 5'-RNA$_{60}$ for 24 hours at 25°C (blue).
**Figure S6.** Time traces of NCLP assembly from NMR. (a) Assembly of NCLPs from P\textsubscript{50}N\textsubscript{405} with polyA-RNA\textsubscript{6}. Assembly was initiated from a sample of P\textsubscript{50}N\textsubscript{405} of 209 µM by adding polyA-RNA\textsubscript{6} to a final concentration of 235 µM. The filled gray circles represent peak intensities measured in individual SOFAST HMQC spectra and summed over the appearing P\textsubscript{50} peaks or the disappearing resonances from N\textsubscript{405}. Red lines correspond to a simultaneous fit of the summed experimental intensities (P\textsubscript{50}: 4, 5, 6, 22, 25, 38, N\textsubscript{405}: 28, 385, 386, 387, 388, 389, 404, 405) to a double-exponential function with common characteristic assembly rates. Apparent assembly rates were 1.9×10\textsuperscript{-3} s\textsuperscript{-1} and 4.0×10\textsuperscript{-4} s\textsuperscript{-1} compared to 2.0×10\textsuperscript{-3} s\textsuperscript{-1} and 3.6×10\textsuperscript{-4} s\textsuperscript{-1} for 5'-RNA\textsubscript{6}. (b) Time course of the assembly of NCLPs with 5'-RNA\textsubscript{60} as monitored by the appearance of P\textsubscript{50} and the disappearance of individual NTD\textsubscript{ARM}, CTD\textsubscript{ARM} and N\textsubscript{TAIL} resonances in the NMR spectra. The filled gray circles represent peak intensities measured in individual SOFAST HMQC spectra, while the red lines correspond to a simultaneous fit of the experimental intensities of all residues shown to a double-exponential function with common characteristic assembly rates but different individual amplitudes. Assembly was initiated from a sample of P\textsubscript{50}N\textsubscript{405} of 202 µM by adding 5'-RNA\textsubscript{60} to a final concentration of 21 µM. The apparent rates for the assembly process in this case were 6.4×10\textsuperscript{-4} s\textsuperscript{-1} and 7.0×10\textsuperscript{-5} s\textsuperscript{-1}. 
Figure S7

Time course of assembly of NCLPs by EM. The assembly of NCLPs was followed in real-time by adding 5’-RNA₆ to P₃₀N₄₅ at time 0 and visualizing this sample by EM at 4, 9, 25 and 67 minutes after mixing the two components.
Figure S8

(a) $^1$H-$^{15}$N HSQC spectra of 53 µM (red) and 7.5 µM (blue) P50N405 with 500 nM 5'-RNA$_{10}$-FAM at the end of the kinetics as measured by fluorescence anisotropy (Figure 4). (b) Electron micrographs of P50N405 (100 µM) capsids formed with 5'-RNA$_{10}$-FAM (32 µM). Scale bars are 100 nm. (c) Steady state $r$ at the end of the kinetics of P50N405 NCLP formation (Figure 4a) reveals a binding curve like dependence.
Figure S9. $^1$H-$^{15}$N correlation spectra of P$_{50N525}$ (blue), N$_{TAIL}$ (green, residues 401-525 of N) and P$_{50N405}$ (red).
Figure S10. (a) Intensity profile of the $^1$H-$^{15}$N HSQC spectrum of P$_{50}$N$_{525}$ after incubation with 5’-RNA$_{60}$ for 24 hours (red). The intensity profile was calculated as $I/I^0$ where $I$ is the intensity of the resonances in the spectrum with RNA and $I^0$ is the intensity in the spectrum of P$_{50}$N$_{525}$ without RNA. A correction was made to the intensity profile (constant offset) to take into account the percentage of protein not assembled into NCLPs under the given experimental conditions (15%). The intensity profile is compared to the previously published intensity profile of purified NCLPs obtained by expression of full-length N in E. coli (blue).[4] (b) EM micrographs of P$_{50}$N$_{525}$ in the presence of 5’-RNA$_{60}$ (top) and 5’-RNA$_{6}$ (bottom). (c) NC assembly kinetics of P$_{50}$N$_{525}$ were acquired from parallel and perpendicular polarised fluorescence light during the time course of 2000 seconds after addition of 500 nM RNA$_{10}$-FAM to various concentrations of P$_{50}$N$_{525}$ (curves from top to bottom: 43 µM, 23 µM, 20 µM, 7.5 µM, 3.8 µM) and fluorescence anisotropy ($r$) was calculated for every time point. All curves were fit with a global fit (red lines) using two or, if necessary, three exponential rates (see Methods). Blue arrow indicates increasing protein concentration. Linearity between the respective rates and the protein concentration was imposed in a global fit. The offset in $r$ was left as a fitting parameter, which is why the fastest rate could only be fit at low P$_{50}$N$_{525}$ concentrations. Resulting kinetic rates are shown in (d). Rates were determined as $(1.95 \pm 0.07) \times 10^3$ M$^{-1}$s$^{-1}$, $(2.44 \pm 0.05) \times 10^2$ M$^{-1}$s$^{-1}$ and $(8.0 \pm 0.2) \times 10^1$ M$^{-1}$s$^{-1}$. 
Figure S11. Dependence of the assembly of NCLPs on the sequence of RNA. P$_{50N_{405}}$ was incubated with different RNAs for 24 hours at room temperature and visualized by negative stain electron microscopy. Assembly of NCLPs is observed for 5'-RNA$_6$ and polyA-RNA$_6$, while no significant assembly is observed for polyU-RNA$_6$. 
Comparison of the $^1$H-$^{15}$N HSQC spectra of P$_{50}$N$_{405}$ in the absence and presence of polyU-RNA$_6$. Spectra are shown of the isolated P$_{50}$N$_{405}$ (red) and the P$_{50}$N$_{405}$ in the presence of a 2-fold molar excess of polyU-RNA$_6$ recorded after incubation at 25°C for 24 hours (blue), showing that no assembly was observed by NMR for polyU-RNA$_6$. 

Figure S12
Comparison of the $^{1}\text{H}^{15}\text{N}$ correlation spectra of $\text{P}_{50}\text{N}_{405}$ in the absence (red) and presence (blue) of 5'-DNA$_6$ (DNA sequence corresponding to 5' RNA$_6$), showing that no assembly was observed upon addition of DNA. Spectra were recorded before addition and 20 hours after addition of the DNA. Protein concentration was 209 µM, DNA concentration 471 µM.
References

[1] J. Marley, M. Lu, C. Bracken, *J. Biomol. NMR* **2001**, *20*, 71–75.
[2] Z. Solyom, M. Schwarten, L. Geist, R. Konrat, D. Willbold, B. Brutscher, *J. Biomol. NMR* **2013**, *55*, 311–321.
[3] P. Schanda, B. Brutscher, *J. Magn. Reson. San Diego Calif 1997* **2006**, *178*, 334–339.
[4] M. R. Jensen, G. Communie, E. A. Ribeiro Jr, N. Martinez, A. Desfosses, L. Salmon, L. Mollica, F. Gabel, M. Jamin, S. Longhi, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 9839–9844.