Capsid Self-Assembly

International Edition: DOI: 10.1002/anie.201602619
German Edition: DOI: 10.1002/ange.201602619

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**

Abstract: *Measles virus RNA genomes are packaged into helical nucleocapsids (NCs), comprising thousands of nucleoproteins (N) that bind the entire genome. N-RNA provides the template for replication and transcription by the viral polymerase and is a promising target for viral inhibition. Elucidation of mechanisms regulating this process has been severely hampered by the inability to controllably assemble NCs. Here, we demonstrate self-organization of N into NC-like particles in vitro upon addition of RNA, providing a simple and versatile tool for investigating assembly. Real-time NMR and fluorescence spectroscopy reveals biphasic assembly kinetics. Remarkably, assembly depends strongly on the RNA-sequence, with the genomic 5′ end and poly-Adenine sequences assembling efficiently, while sequences such as poly-Uracil are incompetent for NC formation. This observation has important consequences for understanding the assembly process.*

Paramyxoviruses are non-segmented negative-strand RNA viruses that express their own machinery for transcription and replication.[1–3] Paramyxoviridae comprise a large number of human pathogens, including measles (MeV), but also the emergent Nipah and Hendra viruses, that present greater than 50% fatality in humans and for which no successful treatment exists. Their genomes are packaged into large helical assemblies, termed nucleocapsids (NC), consisting of a multitude of nucleoprotein (N) copies that bind the entire sequence of the genome.[4–6] The N-RNA complex provides the template for replication and transcription by the viral polymerase complex composed of the phospho- (P) and large (L) proteins.[7] Elucidation of the mechanisms that regulate transcription and replication holds great promise for the development of targeted drugs against paramyxoviral diseases.

[*] Dr. S. Milles,[+] Dr. M. R. Jensen,[+] Dr. G. Communie,[+] D. Maurin, Dr. G. Schoehn, Prof. R. W. H. Ruigrok, Dr. M. Blackledge
Univ. Grenoble Alpes, CNRS, CEA
Institut de Biologie Structurale
Grenoble (France)
E-mail: rob.ruigrok@ibs.fr
martin.blackledge@ibs.fr

[+] These authors contributed equally to this work.

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201602619.

© 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial, and no modifications or adaptations are made.

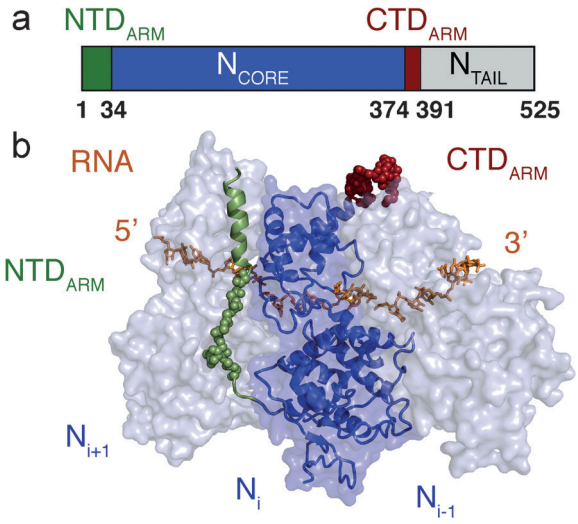

Figure 1. a) Domain organization of MeV N showing the position of the N- and C-terminal arms (NTD_{ARM} and CTD_{ARM}), N_{CORE} and N_{TAIL}. b) Structure of the folded domain of measles virus N in complex with RNA obtained from cryo-electron microscopy on helical NCs (PDB: 4UFT). N_{TAIL} is not shown in this Figure.

MeV N comprises a folded RNA-binding domain (N_{CORE}), as well as an intrinsically disordered C-terminal domain (N_{TAIL}; Figure 1a; Supporting Information, Figure S1).[8,9] A recent cryo-electron microscopy (EM) structure[10] revealed that assembly of NCs is mediated by the NTD_{ARM} and CTD_{ARM} sub-domains, at the N- and C-termini of N_{CORE}, that interact with neighboring N molecules in the NC (Figure 1b), a mode of stabilization also observed in related viruses.[11–14]

The MeV P protein is tetrameric, also containing folded and unfolded domains,[9,15–18] and interacts with N_{CORE} through a 300 amino acid, disordered N-terminal domain, apparently to chaperone RNA-free N prior to NC assembly on the RNA genome,[19] in the so-called N^0P complex. Crystal structures of the N^0P complex of paramyxoviruses[20,21] have been determined by engineering deletion mutants of N_{CORE} devoid of both NTD_{ARM} and CTD_{ARM} sub-domains (as well as N_{TAIL}), to avoid assembly into NC particles. These structures revealed how the N-terminus of P binds to N_{CORE} in a helix–kink–helix conformation. Comparison with the MeV NC[10] revealed that the binding sites of the NTD_{ARM} and CTD_{ARM} in the NCs overlap with binding sites of two α-helices of P in the N^0P complex (Figure 1b), providing a rational explanation of the chaperone activity of P.

Despite this insight into the structural basis of both N^0P and assembled NCs, further understanding of the molecular

mechanisms of NC assembly has been hindered by the inability to isolate intact N^0P and subsequently assemble supramolecular NC particles in vitro. Indeed, assembled NCs from paramyxoviruses could only be purified by over-expressing N in recombinant systems,[4] for example in *E. coli*, where N spontaneously assembles on cellular RNA.

In this study, we engineer a soluble heterodimeric N^0P complex of MeV including the two sub-domains NTD_{ARM} and CTD_{ARM}, and show that the N^0P complex can be triggered to initiate assembly of nucleocapsid-like particles (NCLPs) simply upon addition of RNA. We use nuclear magnetic resonance (NMR), fluorescence spectroscopy, and EM to observe the kinetics of self-assembly of NCLPs and to investigate the molecular basis of NC formation. Remarkably, assembly of MeV NCLPs exhibits strong RNA-sequence dependence. While the 5′ genomic and poly-Adenine sequences form very long NCLPs, certain sequences, such as poly-Uracil, are incompetent for assembly.

An N^0P complex of MeV that is capable of forming NCLPs in vitro was engineered using a fusion construct of MeV P peptide and N, including a TEV cleavage site separating N and P. The minimal peptide required to form the complex was identified using NMR spectroscopy as $P_{1\text{-}50}$ (Figure S2), resulting in a high-yielding, soluble, and stable N^0P complex. This region of P resembles peptides co-crystallized with the N_{CORE} of Nipah virus and MeV (with NTD_{ARM}, CTD_{ARM} and N_{TAIL} deleted).[20,21] RNA-free heterodimeric N^0P complexes of full length N ($P_{50}N_{525}$) and N_{TAIL}-deleted N ($P_{50}N_{405}$) were purified using Ni-affinity chromatography followed by TEV cleavage and gel filtration (Figure S3). MeV NCs obey the "rule of six",[1,2,22] with each N interacting with six nucleotides of the genome. Remarkably, upon addition of $P_{50}N_{405}$ to six-nucleotide RNA strands (corresponding to the 5′ end of the viral genome, 5′-RNA_6 ACCAGA), N spontaneously assembles into NCLPs whose negative-staining EM resembles NCs purified directly by over-expression of N in *E. coli* (Figure 2), demonstrating that continuous RNA is not a prerequisite for formation of NCLPs.

Disordered and flexible regions of $P_{50}N_{405}$ could be observed using solution NMR, corresponding to almost the entire CTD_{ARM} (residues 380–405 of N), the C-terminal end of the P_{50} peptide (residues 44–50 of P), and a few residues belonging to the NTD_{ARM} (residues 26–29; Figure S4). NMR thus can be used to follow NCLP assembly in real time. By adding a molar excess of RNA to $P_{50}N_{405}$ or $P_{50}N_{525}$, and observing the evolution of NMR signals using SOFAST HMQC,[23] a time-resolution of approximately four minutes can be achieved. Resonances of P_{50} appeared as the peptide detaches from the surface of N_{CORE} upon addition of RNA, while flexible resonances of NTD_{ARM} and CTD_{ARM} disappear over time as the NCLPs assemble (Figures 3, S5, S6).

The assembly kinetics of $P_{50}N_{405}$ mixed with 5′-RNA_6 were investigated (Figure 3c). Time traces of N and P were simultaneously fitted using a double exponential function (statistical significance $p < 0.0001$ compared to a single-exponential fit), giving apparent rates $2.0 \times 10^{-3}\ s^{-1}$ and $3.6 \times$

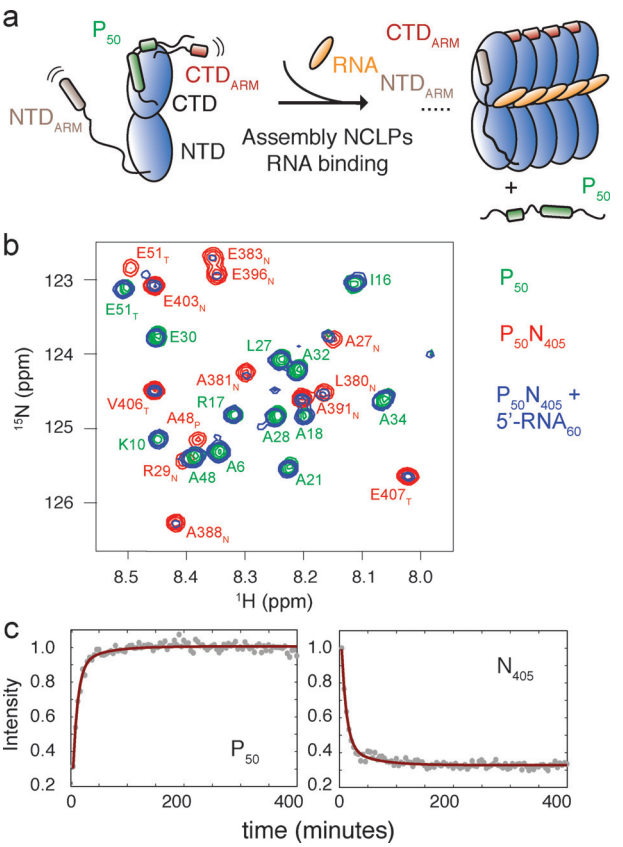

Figure 3. a) Cartoon illustrating the NMR signals that appear or disappear upon assembly of NCLPs. In the N^0P complex, both NTD_{ARM} and CTD_{ARM} of N are flexible, giving rise to observable backbone resonances in ^1H-^{15}N spectra. P_{50} is bound to N so that peaks are too weak to be observed. As N assembles, NTD_{ARM} and CTD_{ARM} stabilize the NCLPs so that these peaks disappear, while resonances from P_{50} grow as more N proteins associate. b) Part of the ^1H-^{15}N spectrum 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). Assignments are shown for the $P_{50}N_{405}$ and P_{50} resonances. Subscript T refers to residues associated with cleavage or affinity tags. c) Assembly of NCLPs from $P_{50}N_{405}$ with 5′-RNA_6 was initiated from 209 μM $P_{50}N_{405}$, adding 5′-RNA_6 to 471 μM. Circles: intensities measured in SOFAST HMQC spectra and summed over appearing P_{50} (left) or disappearing N_{405} (right) resonances (P_{50}: 4, 5, 6, 22, 25, 38, N_{405}: 28, 385, 386, 387, 388, 389, 404, 405), red lines: simultaneous fit to double-exponential with common assembly rates.

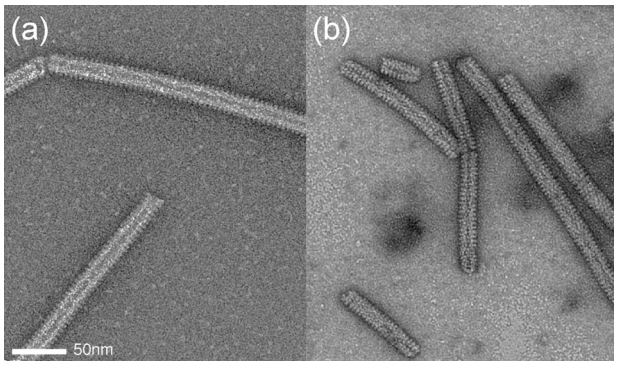

Figure 2. Comparison of a) NCLPs assembled in vitro upon addition of 5′-RNA_6 to $P_{50}N_{405}$ with b) NCLPs obtained from expression of N in *E. coli* and cleavage of N_{TAIL} by trypsin digestion post-purification.

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

$10^{-4}\,s^{-1}$ for 5′-RNA_6 (Figure 3c). Under similar conditions, assembly on 5′-RNA_{60} (corresponding to the first 60 nucleotides of the 5′ end of the viral genome) was found to be slightly slower (Figures 3b and S6b). EM was used to confirm assembly of NCLPs by adding 5′-RNA_6 to $P_{50}N_{405}$ at time 0 and visualizing this sample using EM at increasing time after mixing, showing NCLPs of increasing length (Figure S7).

The absolute assembly rates were determined by measuring time-resolved fluorescence anisotropy of 10-mer RNA with a MeV genome 5′ sequence, labelled at the 3′ terminus with a fluorescein derivative (5′-RNA_{10}-FAM; a longer RNA was used to allow the fluorescein to exit the RNA binding groove and to reduce the potential impact on assembly). The dye attached to free RNA undergoes fast rotation, exhibiting low fluorescence anisotropy, while upon binding to N, NCLP formation and subsequent elongation, the rotational freedom is significantly hindered. Addition of increasing concentrations of $P_{50}N_{405}$ again showed biphasic kinetics (Figure 4) with assembly rates of $(2.2\pm0.2)\times10^3$ and $(3.37\pm0.04)\times10^2\,M^{-1}s^{-1}$. NMR and EM confirmed that the dye did not interfere with NCLP formation (Figure S8).

Both fluorescence and NMR data suggest that an initial reaction dependent on RNA binding is well-separated from the second, slower component, possibly corresponding to assembly of N-coated RNA strands into regular NCLPs. N is thought to bind the RNA immediately as it emerges from the viral polymerase,[3,24] so that RNA_6 may be physiologically relevant. We note that the observed rates of NCLP formation are compatible with the measured elongation speed of the viral polymerase in vivo.[25]

We have previously demonstrated that N_{TAIL} exfiltrates from the inside to the outside of NCLPs through the interstitial space between the N_{CORE} domains where it can interact with the polymerase complex.[8,26,27] The role of N_{TAIL} in assembly was investigated here by measuring NMR, fluorescence, and EM after mixing $P_{50}N_{525}$ and RNA. In addition to backbone resonances that are visible for $P_{50}N_{405}$, $P_{50}N_{525}$ exhibit resonances from the entire N_{TAIL} domain, demonstrating that these residues remain flexible in full-length N (Figure S9). The formation of NCLPs from 5′-RNA_6 and $P_{50}N_{525}$ showed that N_{TAIL} does not interfere with particle assembly in vitro, although the NCLPs are shorter than those produced with $P_{50}N_{405}$ (Figure S10). Notably, a third, significantly slower kinetic rate ($80\pm2\,M^{-1}s^{-1}$) is required to describe the assembly of $P_{50}N_{525}$ compared to $P_{50}N_{405}$, suggesting an additional step in the assembly process involving N_{TAIL}.

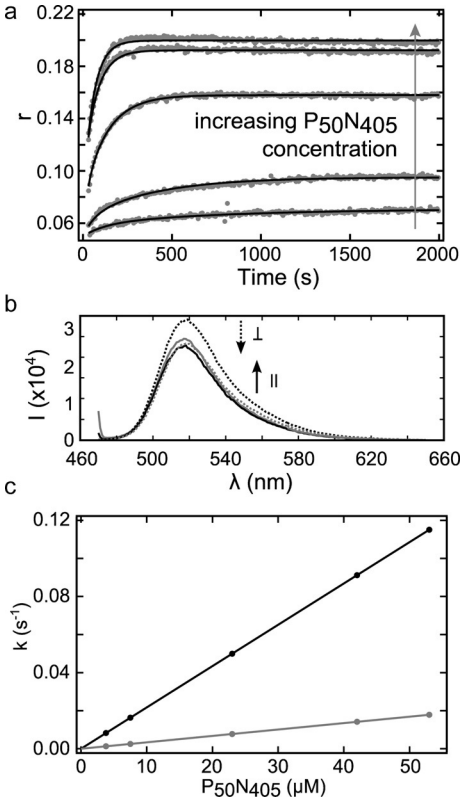

Figure 4. Kinetics of NCLP assembly. a) Parallel and perpendicular polarized fluorescence light was acquired upon addition of 500 nM RNA_{10}-FAM to $P_{50}N_{405}$ (top to bottom: 53, 42, 23, 7.5, 3.8 μM). Fluorescence anisotropy (*r*) was calculated for every time point. All of the curves were fit with a global fit (black lines) using two exponential rates, and imposing linearity between rates and protein concentration (see the Supporting Information). b) Fluorescence spectra at the end of kinetics (dotted lines: perpendicular-polarized; solid line: parallel-polarized fluorescence light) illustrate increased *r* and light scattering at higher protein concentration (black: 7.5 μM; grey: 53 μM $P_{50}N_{405}$). c) Kinetic rates of NCLP formation with $P_{50}N_{405}$ and RNA_{10}-FAM corresponding to (a).

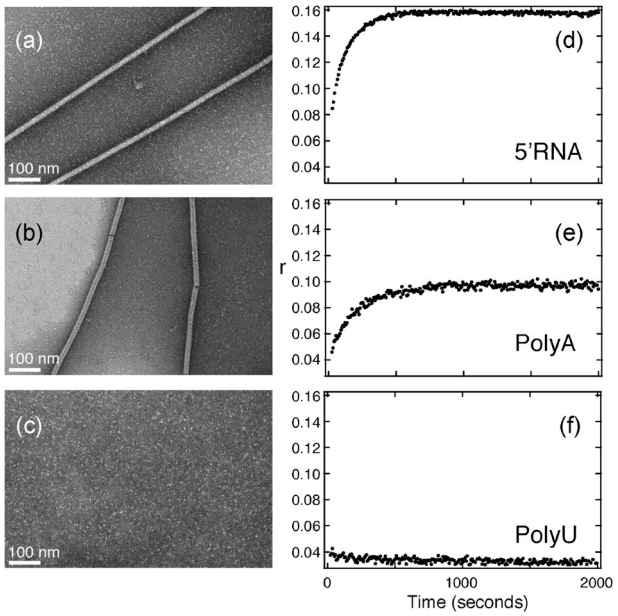

Figure 5. Dependence of NCLP assembly on RNA sequence. (a–c) $P_{50}N_{405}$ was incubated with different RNAs for 24 hours at room temperature and visualized by negative-stain electron microscopy. Assembly of NCLPs is observed for a) 5′-RNA_6, b) polyA-RNA_6, while no assembly is observed for c) polyU-RNA_6 (See Figure S11 for additional representative EMs.) d–f) Fluorescence anisotropy during 2000 seconds after addition of 500 nM, d) 5′-RNA_{10}-FAM, e) $polyA_{10}$-FAM, and f) $polyU_{10}$-FAM to 23 μM $P_{50}N_{405}$.

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Different RNA sequences were tested for their ability to facilitate NCLP assembly in vitro from $P_{50}N_{405}$. Homopolymers comprising purines or pyrimidines alone, UUUUUU (polyU-RNA_6) and AAAAAA (polyA-RNA_6), were compared, as well as genomic 5′-RNA_6. Remarkably, the efficiency of assembly is strongly sequence-dependent (Figures 5 and S11). No assembly is observed for the polyU-RNA_6 samples, while regular NCLPs are obtained for 5′-RNA_6 and polyA-RNA_6 with lengths reaching 1–2 μm. This differential ability to form NCLPs was verified using NMR and fluorescence. No change in fluorescence anisotropy (Figure 5) or displacement of P_{50} peptide over a period of 24 hours at room temperature (Figure S12) were observed upon addition of polyU-RNA_6. RNA-sequence-dependence is surprising considering the apparent necessity of N to encapsidate the entire genome, irrespective of sequence, suggesting the role of an initial specific nucleation site for successful encapsidation. In this respect, we note that the 5′ sequences of both the genome and anti-genome of *paramyxovirinae* show strong conservation of ACCA in the first four nucleotides, followed by positions 5 and 6 occupied mainly by A/G. Indeed, the 5′ sequence of the MeV genome leads to formation of NCs with high efficiency. No assembly is observed on DNA (Figure S13). This specificity may be related to the observed N-RNA interaction in RSV and rabies that is mediated by the 2′ ribose OH group with the carbonyls of N.[12,14]

In all of the structures of paramyxoviral NC particles determined so far, the RNA bases were modelled as generic base-types (uracil), to account for signal averaging over different nucleotides from random cellular RNA. We are therefore unable to determine the molecular mechanisms by which N^0P alone can discriminate between different RNA sequences. In the NC structure determined from cryo-EM,[10] three of the six bases associated with each N stack sequentially and face the surface of the protein, while the following three stack and point into the solvent. The only significant interactions between base moieties and the protein surface occur through a stacking of Y260, and the first of the three stacked-in bases (position 2), with base-specific hydrogen bonding identified between Q202 and the uracil modelled in position 1. Both amino acids are conserved in *paramyxovirinae*. An atomic-resolution structural description of protein–RNA interactions is necessary to fully understand the molecular basis of the observed sequence discrimination.

Based on the rational design of N^0P chaperone mimics, we were able to demonstrate self-assembly of highly regular NCLPs of MeV in vitro, uniquely from protein and RNA. In addition to providing a quantitative determination of assembly rates using real-time NMR and fluorescence spectroscopy, revealing biphasic assembly kinetics, we have demonstrated that assembly occurs both in the presence and absence of the 125 aa disordered N_{TAIL} domain. Importantly, we also identify a strong, and unexpected, dependence of NCLP assembly on the RNA sequence. While the genomic 5′ end and polyA sequences assemble with high efficiency, certain RNA sequences are incompetent for NC formation under our experimental conditions. We note that the ability to controllably assemble stable, protein-based superstructures, whose regularly spaced intrinsically disordered domains can be decorated with sequence-modifiable functionalities, may find diverse biotechnological applications both in vitro and in vivo. Finally, this fluorescence-anisotropy-based MeV NC assembly assay provides a powerful tool for the development of small molecule inhibitors of this essential step in viral replication.

Acknowledgements

This work was funded by the Agence Nationale de la Recherche under ComplexDynamics (MB) and NMRSignal (MRJ) and by Finovi. SM acknowledges the EMBO long-term fellowship (ALTF 468-2014) and EC (EMBOCOFUND2012, GA-2012-600394) through Marie Curie Action. This work used platforms of the Grenoble Instruct Centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). The EM facility is supported by the Rhône-Alpes Region, the Fondation pour la Recherche Medicale (FRM) and FEDER funds.

Keywords: fluorescence spectroscopy · measles · NMR spectroscopy · nucleocapsids · self-assembly

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 9356–9360
Angew. Chem. **2016**, *128*, 9502–9506

[1] E. H. Egelman, S. S. Wu, M. Amrein, A. Portner, G. Murti, *J. Virol.* **1989**, *63*, 2233–2243.
[2] D. Kolakofsky, T. Pelet, D. Garcin, S. Hausmann, J. Curran, L. Roux, *J. Virol.* **1998**, *72*, 891.
[3] R. W. H. Ruigrok, T. Crépin, D. Kolakofsky, *Curr. Opin. Microbiol.* **2011**, *14*, 504–510.
[4] G. Schoehn, M. Mavrakis, A. Albertini, R. Wade, A. Hoenger, R. W. H. Ruigrok, *J. Mol. Biol.* **2004**, *339*, 301–312.
[5] D. Bhella, A. Ralph, R. P. Yeo, *J. Mol. Biol.* **2004**, *340*, 319–331.
[6] A. Desfosses, G. Goret, L. F. Estrozi, R. W. H. Ruigrok, I. Gutsche, *J. Virol.* **2011**, *85*, 1391–1395.
[7] J. Curran, D. Kolakofsky, *Adv. Virus Res.* **1999**, *54*, 403–422.
[8] 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. USA* **2011**, *108*, 9839–9844.
[9] G. Communie, R. W. Ruigrok, M. R. Jensen, M. Blackledge, *Curr. Opin. Virol.* **2014**, *5*, 72–81.
[10] I. Gutsche, A. Desfosses, G. Effantin, W. L. Ling, M. Haupt, R. W. H. Ruigrok, C. Sachse, G. Schoehn, *Science* **2015**, *348*, 704–707.
[11] M. Alayyoubi, G. P. Leser, C. A. Kors, R. A. Lamb, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1792–1799.
[12] R. G. Tawar, S. Duquerroy, C. Vonrhein, P. F. Varela, L. Damier-Piolle, N. Castagné, K. MacLellan, H. Bedouelle, G. Bricogne, D. Bhella, et al., *Science* **2009**, *326*, 1279–1283.
[13] T. J. Green, X. Zhang, G. W. Wertz, M. Luo, *Science* **2006**, *313*, 357–360.
[14] A. A. V. Albertini, A. K. Wernimont, T. Muziol, R. B. G. Ravelli, C. R. Clapier, G. Schoehn, W. Weissenhorn, R. W. H. Ruigrok, *Science* **2006**, *313*, 360–363.
[15] K. Johansson, J.-M. Bourhis, V. Campanacci, C. Cambillau, B. Canard, S. Longhi, *J. Biol. Chem.* **2003**, *278*, 44567–44573.
[16] R. L. Kingston, D. J. Hamel, L. S. Gay, F. W. Dahlquist, B. W. Matthews, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8301–8306.

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[17] P. Bernado, L. Blanchard, P. Timmins, D. Marion, R. W. H. Ruigrok, M. Blackledge, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17002 – 17007.
[18] G. Communie, T. Crépin, D. Maurin, M. R. Jensen, M. Blackledge, R. W. H. Ruigrok, *J. Virol.* **2013**, *87*, 7166 – 7169.
[19] J. Curran, J. Marq, D. Kolakofsky, *J. Virol.* **1995**, *69*, 849 – 855.
[20] F. Yabukarski, P. Lawrence, N. Tarbouriech, J.-M. Bourhis, E. Delaforge, M. R. Jensen, R. W. H. Ruigrok, M. Blackledge, V. Volchkov, M. Jamin, *Nat. Struct. Mol. Biol.* **2014**, *21*, 754 – 759.
[21] S. G. Guryanov, L. Liljeroos, P. Kasaragod, T. Kajander, S. J. Butcher, *J. Virol.* **2015**, *90*, 2849 – 2857.
[22] P. Calain, L. Roux, *J. Virol.* **1993**, *67*, 4822 – 4830.
[23] P. Schanda, B. Brutscher, *J. Am. Chem. Soc.* **2005**, *127*, 8014 – 8015.
[24] O. Gubbay, J. Curran, D. Kolakofsky, *J. Gen. Virol.* **2001**, *82*, 2895 – 2903.
[25] S. Plumet, W. P. Duprex, D. Gerlier, *J. Virol.* **2005**, *79*, 6900 – 6908.
[26] G. Communie, J. Habchi, F. Yabukarski, D. Blocquel, R. Schneider, N. Tarbouriech, N. Papageorgiou, R. W. H. Ruigrok, M. Jamin, M. R. Jensen, et al., *PLoS Pathog.* **2013**, *9*, e1003631.
[27] R. Schneider, D. Maurin, G. Communie, J. Kragelj, D. F. Hansen, R. W. H. Ruigrok, M. R. Jensen, M. Blackledge, *J. Am. Chem. Soc.* **2015**, *137*, 1220 – 1229.

Received: March 15, 2016
Revised: April 22, 2016
Published online: June 6, 2016

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim