Protein Structure

International Edition: DOI: 10.1002/anie.201606609
German Edition: DOI: 10.1002/ange.201606609

Secondary Structure and Membrane Topology of the Full-Length Dengue Virus NS4B in Micelles

*Yan Li, Ying Lei Wong, Michelle Yueqi Lee, Qingxin Li, Qing-Yin Wang, Julien Lescar, Pei-Yong Shi, and CongBao Kang**

Abstract: *Dengue virus nonstructural protein 4B (NS4B) is a membrane protein consisting of 248 residues with a crucial role in virus replication and interference with the host innate immunity. The dengue virus serotype 3 NS4B was reconstituted into lyso-myristoyl phosphatidylglycerol (LMPG) micelles. Backbone resonance assignment of NS4B was obtained using conventional solution NMR experiments. Further studies suggested that NS4B contained eleven helices and six of them form five potential transmembrane regions. This study provides atomic level information for an important drug target to control flavivirus infections.*

Flaviviruses, such as dengue virus, Zika virus, and West Nile virus, are small enveloped viruses. Their viral genome is a single plus-stranded RNA that encodes a polypeptide which is enzymatically processed into three structural proteins (capsid, membrane, and envelope) and seven nonstructural (NS) proteins: NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. All NS proteins play important roles in viral replication and assembly.[1] NS1 is a glycoprotein, playing essential roles in viral RNA synthesis.[2] NS3 and NS5 are soluble proteins with enzymatic activities. The N-terminal region of NS3 containing approximately 180 residues is a serine protease responsible for the maturation of viral polyprotein.[3] The C-terminal region of NS3 harbors both RNA triphosphatase and RNA helicase activities.[4] In addition to its enzymatic activities, NS3 also plays a role in viral assembly.[5] NS5 has both methyltransferase and RNA-dependent RNA polymerase activities.[6] Structural studies for NS1, NS3, and NS5 have been carried out using various methods, such as X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy, which shed light on their functions and led to the development of potent inhibitors.[7]

Four of the flavivirus NS proteins: NS2A, NS2B, NS4A, and NS4B are integral membrane proteins and, compared to the soluble viral proteins, very little is known for them. For the dengue virus, these four membrane proteins play important roles within the viral replication complex. NS2A contains five transmembrane helices and is critical for viral RNA synthesis, viral assembly, and inhibiting the interferon α/β response.[8] The N-terminal fragment of NS2A may be important for production of viral infectious particles.[9] NS2B is a four-span membrane protein.[10] Its main function is to determine the folding and localization of the NS3 protease domain. The cofactor region of NS2B forms a tight non-covalent complex with the NS3 protein, which is required for proper folding of the protease and its activity. A recent study suggested that NS2B can also function as a viroporin.[11] NS4A is a membrane protein with two transmembrane helices.[12] NS4A and NS4B form a complex and can induce membrane rearrangement.[12b,13] Among the NS integral membrane proteins, NS4B is a particularly attractive target to treat dengue and other diseases caused by flaviviruses because several potent inhibitors identified from high throughput screening campaigns using a dengue virus replicon assay have been shown to target NS4B.[14]

Pioneering biochemical studies have shown that NS4B contains three transmembrane helices and the N-terminal region of NS4B contains two membrane interacting helices.[15] Our previous study suggested that the N-terminal two membrane-interacting helices may form transmembrane helices as they are buried in micelles.[16] Unsuccessful attempts to crystallize the NS4B protein suggest that it will be challenging to determine the structure of intact NS4B using X-ray crystallography because NS4B contains a long cytoplasmic region and its N-terminal fragment is dynamic.[13,16] We therefore decided to explore its secondary structure and membrane topology using solution NMR spectroscopy.

The full-length NS4B proteins from all four dengue serotypes can be expressed and purified from *E. coli*. Purified proteins were reconstituted in detergent micelles that are required for both protein purification and folding. NS4Bs exhibited dispersed cross peaks in the ^{1}H-^{15}N-HSQC NMR spectra (Figure S1 in the Supporting Information). Among

[*] Dr. Y. Li, Y. L. Wong, M. Y. Lee, Dr. C. Kang
Experimental Therapeutics Centre, Agency for Science, Technology and Research (A*STAR)
31 Biopolis Way, Nanos, #03-01, Singapore, 138669 (Singapore)
E-mail: cbkang@etc.a-star.edu.sg
Dr. Q. Li
Institute of Chemical & Engineering Sciences, Agency for Science, Technology, and Research
1 Pesek Road, Jurong Island, Singapore 627833 (Singapore)
Dr. Q. Y. Wang
Novartis Institute for Tropical Diseases
Singapore (Singapore)
Dr. J. Lescar
School of Biological Sciences, Nanyang Technological University
Singapore (Singapore)
Dr. P. Y. Shi
Department of Biochemistry & Molecular Biology, Department of Pharmacology & Toxicology, and Sealy Center for Structural Biology & Molecular Biophysics, University of Texas Medical Branch
Galveston, TX (USA)
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under http://dx.doi.org/10.1002/anie.201606609.

these four NS4B proteins, further structural studies were carried out for NS4B from dengue virus serotype 3 in lysomyristoyl phosphatidylglycerol (LMPG) micelles, because of its high yield and increased protein stability. Two strategies were used to prepare samples for NMR study (Figure S2). NS4B was first expressed at low temperature in *E. coli* and bacterial membrane was harvested by ultra-centrifugation. LMPG was then used to extract recombinant protein from the cell membrane and protein was purified by affinity chromatography and gel filtration chromatography. The other purification step was to add LMPG (2%, w/v) in the resuspension buffer during sonication to extract proteins without membrane enrichment. Recombinant protein was then purified as aforementioned. Dimeric NS4B was observed even in SDS-PAGE, when the sample was prepared using the first method (Figure S2) and recombinant protein was mainly in monomeric form when it was purified using the second strategy. As NS4B can form functional dimers during viral replication,[17] enriching *E. coli* cell membrane containing recombinant NS4B through ultra-centrifugation might favor dimer formation. In the second purification method, *E. coli* cell membrane was not enriched and high concentration of detergent micelles may prevent recombinant proteins from forming dimers. NS4B prepared from these two methods exhibited similar ^{1}H-^{15}N-HSQC spectra (Figure S2), which may arise from the fact that the signals from dimeric NS4B could not be observed due to the large molecular weight of the dimer–micelle complex. We used the sample prepared with the second method for NMR studies as the sample was more homogeneous. The first method will be suitable for obtaining a sample with dimeric NS4Bs, but it has been noted that a suitable membrane mimicking system should be chosen because most detergent micelles may not be the idealist system for dimeric protein structural characterization in solution.[18]

NS4B consists of 248 residues and is the largest dengue NS membrane proteins. Assignment of such a protein with helical structures is challenging due to signal overlap and the close values of Cα and Cβ chemical shifts of hydrophobic residues.[19] To obtain backbone assignment, we made another construct containing N-terminal 130 residues of NS4B of dengue 3 (NTD). The ^{1}H-^{15}N-HSQC NMR spectrum of NTD was superimposed with that of the full length (Figure S3), which was used to aid in backbone assignment. Backbone resonance assignments of the full-length NS4B protein were achieved using transverse relaxation-optimized spectroscopy (TROSY)[20]- based 3D heteronuclear experiments based on connections of Cα, Cβ, and C′ carbons (Figure S4). Most backbone amides and amide protons in the ^{1}H-^{15}N-HSQC NMR spectrum were assigned except for residues M1, N2, T50, L156, D157, D162, and C192. 230 peaks out of 236 (excluding 12 proline and 8 residues from the C-terminal LEHHHHHH tag used for protein purification) in the spectrum were assigned (Figure 1A). The assignment has been deposited in the Biological Magnetic Resonance Bank with access number 26846. The secondary structure of NS4B was determined using TALOS+[21] by analyzing the backbone chemical shifts, suggesting that NS4B contains eleven α-helices (Figure 1B). TALOS+ prediction agrees well with the result of chemical shift index[22] by comparing the Cα chemical shifts with the values from random coil structures (Figure S5). Helix α1 is formed by residues L6 to D13, α2 by A38 to E58, α3 by L66 to M77, α4 by P93 to L97, α5 by P105 to Y119, α6 by G125 to K136, α7 by K164 to M185, α8 by E193 to A198, α8′ by P201 to W206, α9 by T217 to R227, and α9′ by Y230 to V244. Among these well identified helical elements, α4 is a short helix. A break between α8 and α8′ is introduced by a proline residue, which was also observed for α9 and α9′ (Figure 1B). The lengths of these helices were further defined by studying the NOE connectivities (Figure S6). Based on both TALOS+ prediction and NOE analysis, NS4B in LMPG micelles contains eleven helices: α1 (E3–D13), α2 (S37–S60), α3 (T61–G82), α4 (P93–Y100), α5 (P105–G123), α6 (G125–R137), α7 (S163–M185), α8 (E193–G200), α8′ (P201–E207), α9 (T217–R227), and α9′ (Y230–V244). The precise lengths and boundaries of these identified helices are slightly different. This may be due to different characteristics and exposures such as solvent accessibility, membrane interactions, and transmembrane. The secondary structures of first 125 residues in full-length NS4B were similar to those of a construct from the previous study in which only these residues of dengue 2 NS4B were purified and characterized.[16] The long cytoplasmic loop (G125–D162) was shown to be important for NS3 binding.[13] Previous NMR study on this loop suggested that there may be a short helix formed by residues Q127 to Q135.[13] This region was also shown to be a helix (part of α6) in this study, supporting the previous findings. Helices α8 and α8′ are separated by a proline residue, suggesting that these two helices may form a single kinked helix. Likewise, helices α9 and α9′ may form a kinked transmembrane helix and G228 and S229 form a kink in this helix. NS4B contains several proline residues that can serve as a helix breaker. There is a proline in α2 and this residue is conserved in the NS4Bs from the flavivirus.[23] The proline (P201) between α8 and α8′ is conserved among the four dengue serotypes, not in other flaviviruses such as West Nile virus and Zika virus, which suggests that the structure of this helix might be different between dengue NS4Bs and other flaviviruses (Figure S7).[23] Studies have also shown that some proline residues are critical for NS4B function. For instance, the mutation of proline in dengue 4 NS4B (P104/L) can abolish its interaction with NS3 protein.[24] These helix breakers in NS4B might have important functions by altering the continuity of a helix.

NS4B contains three cysteine residues, namely, C99, C178, and C192. Only C99 is conserved in flavivirus and the other two are conserved among dengue four serotypes (Figure S7). Cβ chemical shifts of C99 and C178 were $\delta = 40.95$ and 40.17 ppm, respectively. The redox states of these two residues are oxidized as the Cβ chemical shifts are above 35 ppm, suggesting that these two residues might form a disulfide bond under such conditions, which might be important for NS4B function if this bond exists.[25] C192 was not unambiguously assigned, which may be due to its dynamics or exchanges.

The dynamics of NS4B were analyzed by obtaining ^{15}N T_1, T_2, and heteronuclear NOEs (hetNOEs; Figure 2). It has been noted that most residues in the transmembrane regions could

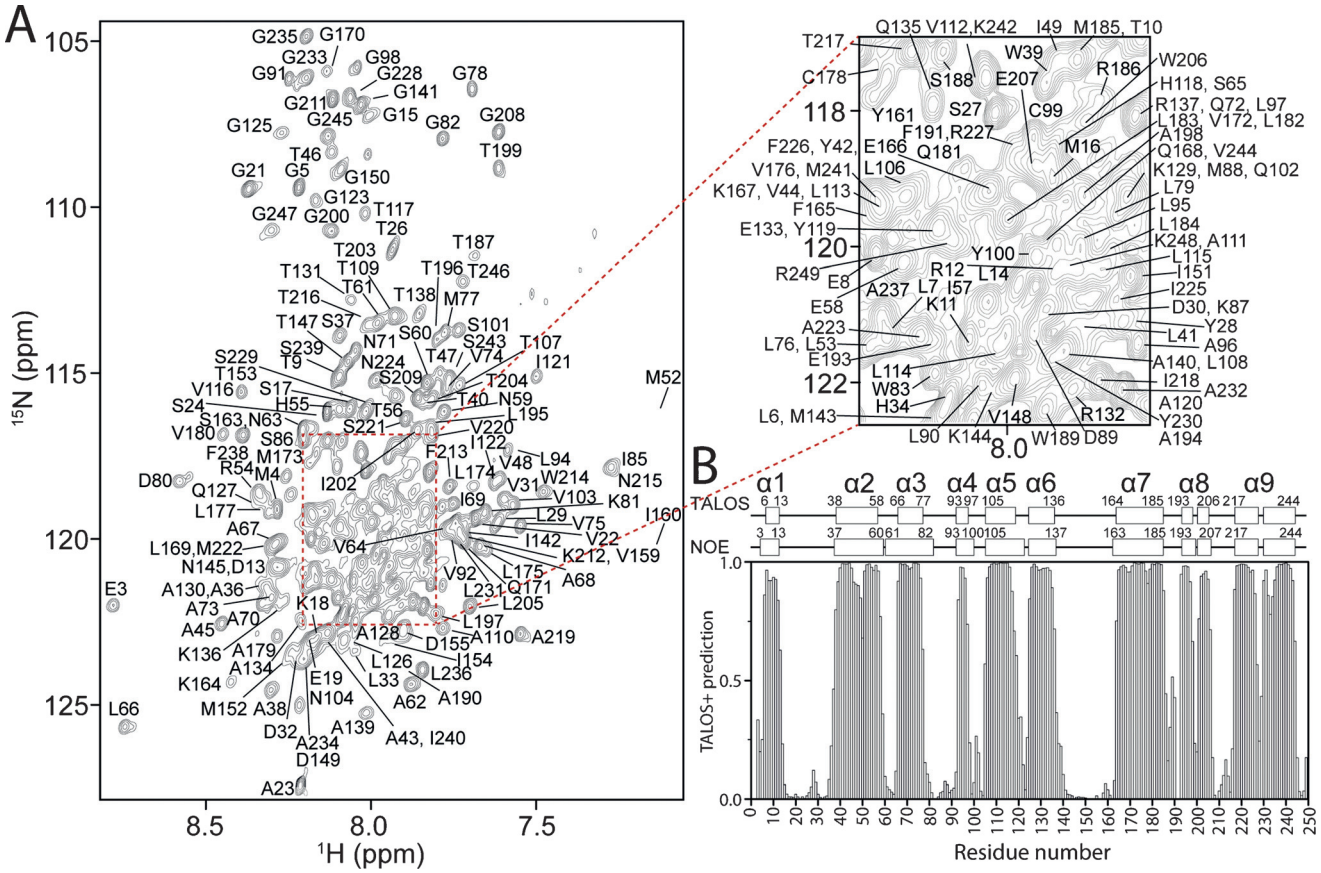

Figure 1. A) Assignment of dengue 3 NS4B in LMPG micelles. ^{1}H–^{15}N-TROSY NMR spectrum of NS4B was collected at 313 K. The assignment of each peak is labeled with residue name and sequence number. The middle region is shown on the right side. B) Secondary structure analysis using TALOS and NOE connections. TALOS+ analysis was carried out using an input file containing HN, N, Cα, Cβ, and C′ chemical shifts. The possibility is plotted against residue number. The possible to be a helix is ranked from 0 (not a helix) to 1 (helix). The helix and coil structures are shown as boxes and lines, respectively.

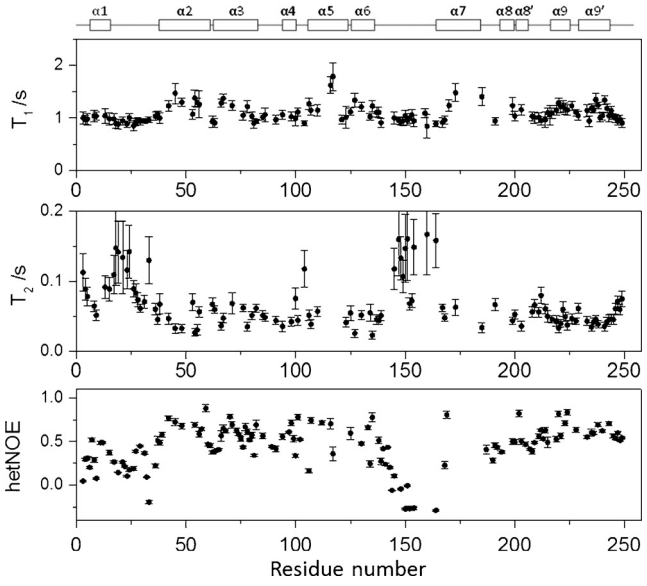

Figure 2. ^{15}N T_1, T_2, and hetNOE relaxation data for NS4B. The data was collected at 700 MHz. Only residues that can be analyzed are shown.

not be analyzed due to signal overlaps, which is not surprising for a membrane protein with such size. Based on the analysis of residues that have no signal overlap, the relaxation results are still helpful to understand the dynamic nature of NS4B in micelles. It is clear that the first 30 residues including α1 are dynamic, which is supported by the low hetNOE values (less than 0.6). The longest cytoplasmic loop region containing G125 to D162 was shown to be highly dynamic in micelles, which is characterized by the low hetNOE values when it was expressed and purified alone in solution.[13] In the full-length NS4B, only the C-terminal region of the loop exhibited low and negative hetNOEs, suggesting high mobility. The N-terminal region of the loop region containing α6 was more stable than the C-terminus, suggesting that micelle or membrane environment may stabilize the helical structures. Residues from helices α2, α3, α5, α7, α8, and α9 exhibited higher T_1 than those from other regions.

To understand the membrane topology of NS4B in LMPG micelles, we carried out a paramagnetic relaxation enhancement (PRE) experiment. The solvent accessibility of NS4B was explored by comparison of the peak intensities of a residue in the absence and presence of gadolinium (Figure S8). Residues from 142 to 160 and residues from the N-

terminus exhibited significant changes in peak intensities upon addition of gadolinium, suggesting that these residues are exposed to the solvent. Residues from 5 to 30, 125 to 140 exhibited medium changes, suggesting that they were interacting with micelles. Other residues were shown to be protected from exposure to gadolinium, suggesting that they may be transmembrane regions or buried in the micelles. We also used hydrogen–deuterium (H–D) exchange experiments to further understand the structure and membrane topology of NS4B (Figure S9). When NS4B was exposed to D_2O, most peaks in the ^1H-^{15}N-HSQC spectrum disappeared except for residues from α2, α3, α5, α7, and α9. Most protected residues were hydrophobic ones such as Leu, Ala, and Val. These residues were protected from H–D exchanges by forming a secondary structure or/and being buried deeply in detergent micelles. A biochemical study suggested that α5 and α7 are transmembrane helices.[15] Several residues from these two helices were protected from exchanges, confirming that they are transmembrane helices. Several residues from α2 and α3 were also protected from exchanges with faster exchange rates than those of residues from α5 and α7. Only one residue from α9 and α9′ exhibited a cross peak in the ^1H-^{15}N-HSQC spectrum, suggesting that these helices are different from the other four helices (Figure S9). Based on our PRE and H–D exchange experiments, NS4B might contain five transmembrane helices-α2, α3, α5, and α7 (Figure 3A). Helices α9 and α9′ should be a transmembrane domain with a break in viral poly protein because NS5 protein is present in the cytoplasm (Figure 3B). These two helices were protected in the PRE experiment (Figure S8), suggesting that it may be a transmembrane helix. H–D exchange experiment demonstrated that α9 and α9′ were not protected from exchanges, indicating that these helices are not behaving similarly to other transmembrane regions in vitro, suggesting that this transmembrane domain may undergo conformational changes when NS5 is released from NS4B (Figure 3C). This result is also consistent with the previous finding that this region may flip over the cell membrane.[15]

In summary, we conducted backbone resonance assignments of full-length NS4B in LMPG micelles. Secondary structures of NS4B were obtained based on analysis of the obtained chemical shifts and NOE connections. NS4B might contain five transmembrane helices based on H–D exchange and PRE experiments. The last transmembrane helix which contains a break, is membrane interacting, and is buried in micelles. Residues from 142 to 160 were found to be exposed to the solvent, which is consistent with the finding that they form part of the cytoplasmic loop that interacts with the NS3 protein.[24] Overall this study provides, to our knowledge, the most detailed structural description of NS4B that will be of great use to develop specific inhibitors against flaviviruses.

Acknowledgements

C.K. appreciates support by A*STAR BMRC IAF grant (111105) and A*STAR JCO grant (1431AFG102/1331A028). J.L. acknowledges support from MOH by grant NMRC/

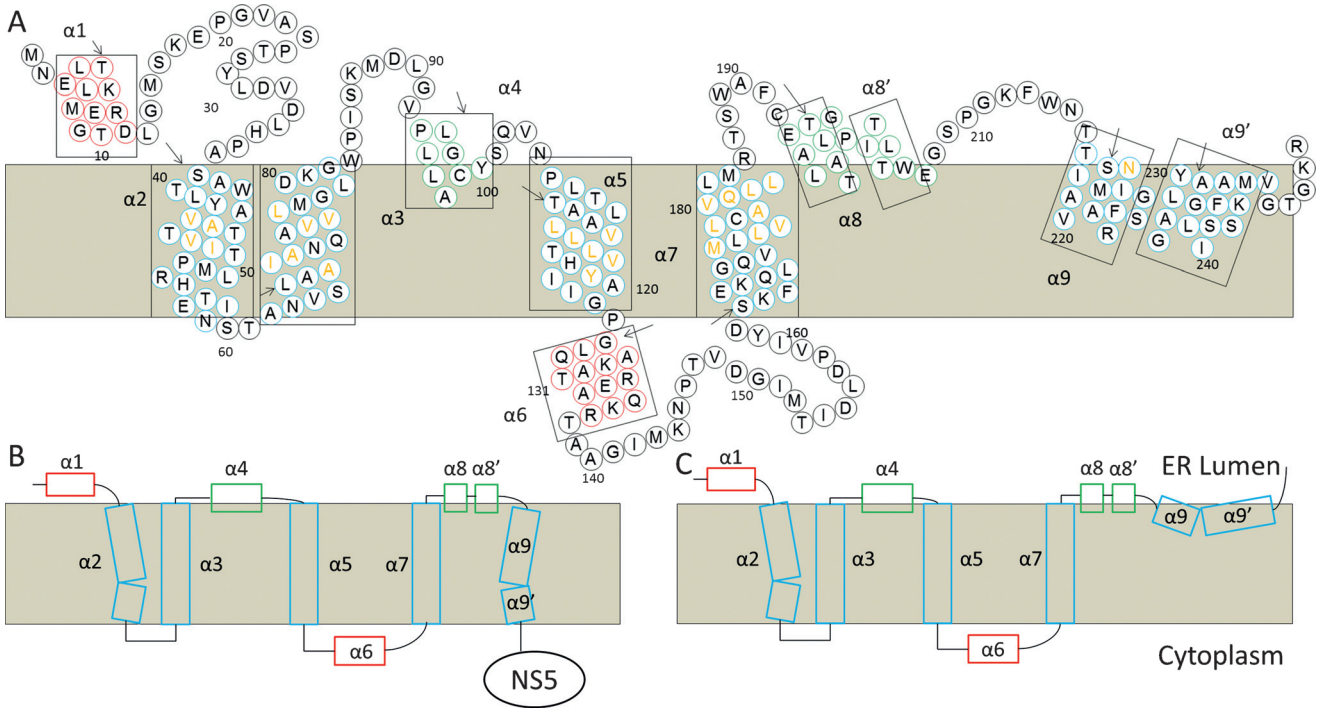

Figure 3. Membrane topology of NS4B. A) NS4B sequence and its topology are predicted based on PRE and H–D exchange experiments. Water-soluble residues are highlighted in red, membrane interacting residues are in green, and residues protected from exposure to gadolinium blue. Residues highlighted in brown are protected from H–D exchanges. Helices α9 and α9′ behave differently compared to other transmembrane helices, suggesting that they may be buried in the micelles. M1 is an artificial residue for protein expression in *E. coli*. B),C) Simplified model of NS4B topology in the presence (B) and absence (C) of NS5.

CBRG/0073/2014. Q.L. thanks A*STAR the Biospecialties project (1526004161). We thank colleagues from Novartis for NMR data collection.

Keywords: dengue virus · membrane protein · membrane topology · NMR spectroscopy · NS4B

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 12068–12072
Angew. Chem. **2016**, *128*, 12247–12251

[1] P. D. Uchil, V. Satchidanandam, *J. Biol. Chem.* **2003**, *278*, 24388–24398.
[2] T. P. Wallis, C. Y. Huang, S. B. Nimkar, P. R. Young, J. J. Gorman, *J. Biol. Chem.* **2004**, *279*, 20729–20741.
[3] a) B. Falgout, M. Pethel, Y. M. Zhang, C. J. Lai, *J. Virol.* **1991**, *65*, 2467–2475; b) C. F. Arias, F. Preugschat, J. H. Strauss, *Virology* **1993**, *193*, 888–899.
[4] a) D. Luo, S. G. Vasudevan, J. Lescar, *Antiviral Res.* **2015**, *118*, 148–158; b) T. Xu, A. Sampath, A. Chao, D. Wen, M. Nanao, P. Chene, S. G. Vasudevan, J. Lescar, *J. Virol.* **2005**, *79*, 10278–10288.
[5] C. G. Patkar, R. J. Kuhn, *J. Virol.* **2008**, *82*, 3342–3352.
[6] a) D. Ray, A. Shah, M. Tilgner, Y. Guo, Y. Zhao, H. Dong, T. S. Deas, Y. Zhou, H. Li, P. Y. Shi, *J. Virol.* **2006**, *80*, 8362–8370; b) M. Ackermann, R. Padmanabhan, *J. Biol. Chem.* **2001**, *276*, 39926–39937.
[7] a) D. Luo, T. Xu, C. Hunke, G. Gruber, S. G. Vasudevan, J. Lescar, *J. Virol.* **2008**, *82*, 173–183; b) C. G. Noble, C. C. Seh, A. T. Chao, P. Y. Shi, *J. Virol.* **2012**, *86*, 438–446; c) L. Ma, C. T. Jones, T. D. Groesch, R. J. Kuhn, C. B. Post, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3414–3419; d) E. Pokidysheva, Y. Zhang, A. J. Battisti, C. M. Bator-Kelly, P. R. Chipman, C. Xiao, G. G. Gregorio, W. A. Hendrickson, R. J. Kuhn, M. G. Rossmann, *Cell* **2006**, *124*, 485–493; e) L. de la Cruz, T. H. Nguyen, K. Ozawa, J. Shin, B. Graham, T. Huber, G. Otting, *J. Am. Chem. Soc.* **2011**, *133*, 19205–19215; f) Y. M. Kim, S. Gayen, C. Kang, J. Joy, Q. Huang, A. S. Chen, J. L. Wee, M. J. Ang, H. A. Lim, A. W. Hung, R. Li, C. G. Noble, T. Lee le, A. Yip, Q. Y. Wang, C. S. Chia, J. Hill, P. Y. Shi, T. H. Keller, *J. Biol. Chem.* **2013**, *288*, 12891–12900.
[8] a) X. Xie, S. Gayen, C. Kang, Z. Yuan, P. Y. Shi, *J. Virol.* **2013**, *87*, 4609–4622; b) B. M. Kummerer, C. M. Rice, *J. Virol.* **2002**, *76*, 4773–4784; c) J. Y. Leung, G. P. Pijlman, N. Kondratieva, J. Hyde, J. M. Mackenzie, A. A. Khromykh, *J. Virol.* **2008**, *82*, 4731–4741.
[9] S. Vossmann, J. Wieseler, R. Kerber, B. M. Kummerer, *J. Virol.* **2015**, *89*, 4951–4965.
[10] a) S. Clum, K. E. Ebner, R. Padmanabhan, *J. Biol. Chem.* **1997**, *272*, 30715–30723; b) Y. Li, Q. Li, Y. L. Wong, L. S. Liew, C. Kang, *Biochim. Biophys. Acta Biomembr.* **2015**, *1848*, 2244–2252.
[11] M. León-Juárez, M. Martínez-Castillo, G. Shrivastava, J. García-Cordero, N. Villegas-Sepulveda, M. Mondragón-Castelán, R. Mondragón-Flores, L. Cedillo-Barrón, *Virol. J.* **2016**, *13*, 1–11.
[12] a) M. H. Lin, H. J. Hsu, R. Bartenschlager, W. B. Fischer, *J. Biomol. Struct. Dyn.* **2014**, *32*, 1552–1562; b) S. Miller, S. Kastner, J. Krijnse-Locker, S. Buhler, R. Bartenschlager, *J. Biol. Chem.* **2007**, *282*, 8873–8882.
[13] J. Zou, T. Lee le, Q. Y. Wang, X. Xie, S. Lu, Y. H. Yau, Z. Yuan, S. Geifman Shochat, C. Kang, J. Lescar, P. Y. Shi, *J. Virol.* **2015**, *89*, 3471–3483.
[14] a) Q. Y. Wang, H. Dong, B. Zou, R. Karuna, K. F. Wan, J. Zou, A. Susila, A. Yip, C. Shan, K. L. Yeo, H. Xu, M. Ding, W. L. Chan, F. Gu, P. G. Seah, W. Liu, S. B. Lakshminarayana, C. Kang, J. Lescar, F. Blasco, P. W. Smith, P. Y. Shi, *J. Virol.* **2015**, *89*, 8233–8244; b) K. W. van Cleef, G. J. Overheul, M. C. Thomassen, J. M. Marjakangas, R. P. van Rij, *Antimicrob. Agents Chemother.* **2016**, *60*, 2554–2557.
[15] S. Miller, S. Sparacio, R. Bartenschlager, *J. Biol. Chem.* **2006**, *281*, 8854–8863.
[16] Y. Li, Y. M. Kim, J. Zou, Q. Y. Wang, S. Gayen, Y. L. Wong, T. Lee le, X. Xie, Q. Huang, J. Lescar, P. Y. Shi, C. Kang, *Biochim. Biophys. Acta Biomembr.* **2015**, *1848*, 3150–3157.
[17] J. Zou, X. Xie, T. Lee le, R. Chandrasekaran, A. Reynaud, L. Yap, Q. Y. Wang, H. Dong, C. Kang, Z. Yuan, J. Lescar, P. Y. Shi, *J. Virol.* **2014**, *88*, 3379–3391.
[18] M. Zhang, R. Huang, R. Ackermann, S. C. Im, L. Waskell, A. Schwendeman, A. Ramamoorthy, *Angew. Chem. Int. Ed.* **2016**, *55*, 4497–4499; *Angew. Chem.* **2016**, *128*, 4573–4575.
[19] a) C. Huang, S. Mohanty, *J. Am. Chem. Soc.* **2010**, *132*, 3662–3663; b) A. Gautier, J. P. Kirkpatrick, D. Nietlispach, *Angew. Chem. Int. Ed.* **2008**, *47*, 7297–7300; *Angew. Chem.* **2008**, *120*, 7407–7410.
[20] a) K. Pervushin, R. Riek, G. Wider, K. Wuthrich, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366–12371; b) M. Salzmann, K. Pervushin, G. Wider, H. Senn, K. Wuthrich, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13585–13590.
[21] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, *J. Biomol. NMR* **2009**, *44*, 213–223.
[22] D. S. Wishart, B. D. Sykes, F. M. Richards, *Biochemistry* **1992**, *31*, 1647–1651.
[23] J. A. Wicker, M. C. Whiteman, D. W. Beasley, C. T. Davis, C. E. McGee, J. C. Lee, S. Higgs, R. M. Kinney, C. Y. Huang, A. D. Barrett, *Virology* **2012**, *426*, 22–33.
[24] I. Umareddy, A. Chao, A. Sampath, F. Gu, S. G. Vasudevan, *J. Gen. Virol.* **2006**, *87*, 2605–2614.
[25] D. Sharma, K. Rajarathnam, *J. Biomol. NMR* **2000**, *18*, 165–171.

Received: July 7, 2016
Published online: August 24, 2016

