PEARLS
Capsid protein is central to the birth of flavivirus particles

Ter Yong Tan1,2, Guntur Fibriansah1,2, Shee-Mei Lok1,2*

1 Programme in Emerging Infectious Diseases, Duke–National University of Singapore Medical School, Singapore, Singapore, 2 Centre for BioImaging Sciences, Department of Biological Sciences, National University of Singapore, Singapore

* sheemei.lok@duke-nus.edu.sg

Overview of flavivirus infection cycle

Flaviviruses such as dengue virus (DENV) and Zika virus (ZIKV) are enveloped, positive-sense single-stranded RNA viruses. Upon entry into a host cell by receptor-mediated endocytosis, the virus particle undergoes low-pH–driven endosomal membrane fusion to release its genome into the cytoplasm [1]. This viral genome is then translated into a single polyprotein that is co- and post-translationally processed into 3 structural and 7 nonstructural proteins. The 3 structural proteins that form the immature virus particle are the Capsid (C), Envelope (E), and precursor Membrane (prM) proteins. After formation of the core, consisting of the C protein complexed with the newly synthesized viral RNA genome, the core then buds into the endoplasmic reticulum (ER) membrane decorated with membrane-anchored viral E and prM proteins, forming an immature virus (Fig 1A). The immature virion then undergoes maturation during transport through the secretory pathway (Fig 1B). This involves large conformational rearrangements of the E proteins (Fig 2A), which enhances the cleavage of the prM to M protein by furin proteases (Fig 1C) to generate the infectious mature virus particle (Fig 1D).

The surface protein structures prM and E of the immature and mature flavivirus particles

Cryo-electron microscopy (cryoEM) structures of the immature and mature flavivirus show that their surface proteins, although they are organized in a vastly different way, both exhibit icosahedral symmetries (Fig 2A) [2–4]. These virus surfaces are made up of 180 copies each of the E and prM/M (prM in immature and M in mature) proteins [2]. On the surface of the immature virus in the neutral pH environment of the ER where they are first assembled, they contained 60 spikes (Fig 2A, leftmost). This surface architecture is made from an assembly of 60 inverted tripods that crisscross with each other. Each of these inverted tripods consists of 3 prM–E heterocomplexes (Fig 2B). The base of the inverted tripod (Fig 2B) is anchored to the underlying lipid membrane by their transmembrane (TM) helices. Each of the ectodomains of these 3 prM–E complexes then interacts with those from other neighboring tripods, imparting a spiky appearance to the immature virus (Fig 2A, leftmost).

During the egress of the immature virus through the acidic compartments of the trans-Golgi network, the low pH triggers the reorganization of the heterotrimeric spikes into a dimeric surface protein organization (Figs 1A–1C and 2A) [5]. This dramatic structural rearrangement of the surface glycoprotein is a prerequisite for exposing the furin-cleavage site on the prM for processing by the furin protease (Fig 1C). Once the prM is cleaved, the pr peptide remains bound to the virus at an acidic pH and only detaches upon excretion of the particle into the neutral pH extracellular environment (Fig 1C and 1D) [5]. The final infectious mature
flavivirus particles have a smooth spherical appearance, and the E proteins exist as homodimers lying flat on the surface of the virus lipid membrane (Figs 1D, 1F and 2A) [2, 6, 7]. Three such E protein dimers are laying parallel to each other, forming a raft, and 30 such rafts are arranged in a herringbone pattern (Fig 2A, rightmost).

**Flavivirus C proteins are instrumental in gluing the virus RNA genome and surface proteins together for the formation of infectious virus particle**

Although the virus assembly process is a crucial stage in the virus infection cycle (Fig 1A), its molecular mechanism is poorly understood. C protein, the first flaviviral protein being translated, plays a central role in interacting with the viral RNA genome for packaging into the virus particle during the assembly process. However, how the C protein–RNA complex can associate with the ER lipid membrane that has prM/E proteins anchored on it for virus budding is largely unknown.

While translating the virus RNA genome, the synthesized viral polyprotein is threaded back and forth through the ER membrane; while some of the viral proteins are facing the cytoplasm, the others face the lumen of the ER. The full-length C (f-C) protein, at this point of processing, is largely facing the cytoplasm. The f-C protein contains a disordered N-terminus followed by 5 α-helices (helices α1–α5). The f-C protein helix α5 is a short hydrophobic helix that immediately precedes the prM in the polyprotein, and it helps the prM traverse the ER membrane so as to bring the prM into the ER lumen. The N- and C-terminal ends of this helix α5 in the ER membrane contain cleavage sites for viral NS2B/NS3 protease and host signal peptidase, respectively. When fully processed by these 2 proteases, the end result is the mature C (m-C) protein containing helices α1–α4 in the cytoplasmic side of the ER, leaving behind the helix α5 in the ER membrane, and the prM protein in the ER lumen.

NMR and X-ray crystal structures of the flavivirus m-C dimer show there is structural conservation among the different flaviviruses despite their poor sequence similarity [8–11] (Fig 2C). The 4 α-helices of each protomer are held together by short loops, and they form a 3-layer structure. There is an asymmetric charge distribution on the C dimeric surface: one side contains hydrophobic patches that are mainly contributed by helices α2, while the opposite side contains positively charged residues of helices α4. This leads to the postulation that the hydrophobic residues on one side likely interact with the viral lipid membrane, while the α4 helices on the opposite side interact with the negatively charged RNA genome [10].

Unlike other flaviviruses, the ZIKV m-C protein has a long pre-α1 loop that precedes a short helix α1 [11]. The helix α1 in ZIKV is shorter compared to the other flaviviruses because
there are 2 proline residues, thus breaking the helix. The resulting long pre-α1 loop extends the dimerization surface and enhances interactions between the 2 m-C protomers [11]. The orientation of helix α1 with respect to helix α2 can affect the accessibility of the hydrophobic residues on helix α2. Interestingly, the helix α1 of neuropathic flaviviruses such as ZIKV and West Nile virus (WNV) is positioned perpendicular to helix α2 in a “close conformation” [12].
Non-neuropathic flavivirus like DENV have helix α1 lying parallel to helix α2 in an “open conformation” [10]. The open conformation allows more hydrophobic residues from helix α2 to be more accessible; however, its functional significance is unknown.

Even though the cryoEM structures of surface proteins on mature and immature flavivirus particles (WNV, DENV, ZIKV, and yellow fever virus) have been determined, some to near-atomic and subnanometer resolutions, the densities of the C proteins inside these virus particles are either poor or absent. The first glimpse of a low-resolution density, sandwiched between the core and viral lipid membrane, that somewhat correlates to the volume of the C protein dimer was reported in the immature ZIKV (ImmZIKV) by Prasad and colleagues [13]. This suggests that ImmZIKV could have a more stable C protein structure. Our recent paper, Tan and colleagues [14], uses an antigen-binding fragment (Fab) of HMAb DV62.5, which binds across the prM and E proteins, to stabilize the ImmZIKV surface, which in turn stabilizes the entire virus particle, allows us to observe the C dimeric protein density to 9-Å resolution (Fig 2D). We then fitted the crystal structure of ZIKV m-C protein into this density and showed that the C dimers interact with both the RNA core and the E–prM TM region (Fig 2B). This suggests the C dimer may be the factor that ensures all virus particle are packaged with the RNA genome.

We also observed in the cryoEM density map that some of the C proteins inside the virus particle are actually that of the f-C protein [15], showing clear density of the helix α5, although it is slightly weaker compared to the rest of the C protein. Consistent with the cryoEM density map, mass spectrometry and SDS-PAGE of purified virus sample showed a mixture of m-C and f-C proteins. We therefore fitted helix α5 into its corresponding density at the C-terminal end of the fitted m-C crystal structure and showed that helix α5 plays a big role in the organization of the overall quaternary ImmZIKV structure. The neighboring α5 helices from adjacent f-C dimers interact with each other via hydrophobic interactions to form trimers of f-C dimers (Fig 2E).

The results showed the basic building block that makes up the final architecture of the virus particle. They are the trimers of a mixture of f- or m-C dimers complexed with 9 preformed prM–E proteins (Fig 2F). The ectodomains of prM–E complexes of these basic building blocks then interact with each other on the lumen side of the ER, assembling into an icosahedral virus particle (Fig 2G).

**Can flavivirus assemble in the absence of C protein?**

Coexpression of only the prM and E from flaviviruses such as tick-borne encephalitis virus and DENV showed that these proteins can rapidly assemble into empty subviral particles, and the organization of these proteins are very different from that of the infectious virus particles [16]. The subviral particles showed different sizes, and the most abundant form contains only 60 prM/E proteins on its surface, in contrast to the 180 copies of prM/E on the infectious virus particles [16]. This observation suggests that lateral interactions between prM–E complexes could somewhat partially drive virus assembly, but additional factors are needed for the formation of the correct infectious virus architecture.

**What happens to the C protein within the virus particle during flavivirus maturation?**

Unlike the immature virus (Fig 1E), cryoEM density maps of the mature form of flavivirus do not show any C protein density (Fig 1F). Previously, we have determined the structural rearrangements of the surface proteins prM and E during the DENV maturation process, and it shows that the cluster of 3 prM and E TM regions in the immature virus reorganized into a
dimeric structure in the mature virion [2]. In our current paper, since a C protein dimer interacts simultaneously with the cluster of E/prM TM regions in the immature virus, the dramatic reorganization of these TM regions during maturation may cause C protein to dislodge from prM and E proteins. We speculate that this may cause the C-RNA to condense to the inside the core of the mature virus, thus allowing the release of the C-RNA core into the cytoplasm after the virus and endosomal membrane have fused in the next cycle of infection.

**Comparison of the flavivirus assembly process with other viruses**

Alphaviruses are small, enveloped, positive-sense single-stranded RNA arboviruses that are closely related to flavivirus. CryoEM structures of these viruses show that both follow icosahedral symmetry, although alphavirus has a T = 4 whereas flavivirus has T = 3 symmetry. They have structurally similar surface glycoproteins—E1 in alphavirus and E protein in flavivirus—and an inner C protein layer that surrounds an RNA core. However, their assembly processes are very different—alphavirus is assembled on the plasma membrane with a preassembled nucleocapsid core (RNA–C protein complex), while in flavivirus, virus assembles in the ER in a concerted manner along with the translation of the polypeptide chains [14, 17–19]. It is possible that the flavivirus assembly process may somewhat be more applicable to other envelope viruses that bud off from the ER. One such virus is hepatitis B virus (HBV); however, empty particles with no DNA genome form the majority of the particles released from the cell, which indicates that its DNA-binding protein (equivalent with C protein in flavivirus) may not engage directly with its surface protein, and hence, its assembly process is likely very different [20]. Coronavirus, an RNA envelope virus, also assembles in the ER, with the binding of surface protein M to the N protein–RNA complex packaging the genome into the viral particle [21]. However, the viral particle is pleomorphic (not organized in icosahedral symmetry), suggesting that there is no single assembly unit (a brick in the house) to form the final virus architecture, unlike flaviviruses [22]. This shows that the mechanism of virus assembly process can be quite diverse.

**Can we therapeutically target C protein to interfere with flavivirus assembly?**

The feasibility of targeting C protein has been demonstrated in a number of viruses such as HBV and HIV and even flaviviruses such as DENV [23, 24]. The small-molecule drug ST-148, for instance, has been shown to induce DENV C protein self-association/aggregation [25, 26]. This effect of ST-148 is thought to interfere with virus nucleocapsid assembly, resulting in a broad antiviral activity against all serotypes of DENV that is seen both in vitro and in vivo. Peptide mimetic drugs such as pep14-23, which mimics the N-terminus of DENV C protein, have been shown to block C interaction with lipid droplets (LDs) [27]. Because perturbing C–LD interaction has been shown to reduce virus production, a peptide-based drug that targets other functions of C is another promising way of reducing viremia.

Our study provides a structural understanding of C protein organization inside the immZIKV particle and illuminates its importance in the flavivirus assembly process. Our result may provide targets for designing drugs that could either prevent C protein from interacting with the RNA genome or the TM region of E and prM.

**Conclusion**

The current cryoEM studies show that the C protein is important for virus assembly; its roles include (1) recruiting the RNA genome and (2) interacting with the transmembrane regions of prM–E, and also, (3) the presence of some f-C proteins may determine the spacing between
the prM–E proteins, thus making the correct T = 3 infectious flaviviruses particle conformation. Taken together with the greater appreciation of C in the flavivirus assembly process, we hope that this will spur the development of therapeutics that target virus assembly.

References

1. Modis Y, Ogata S, Clements D, Harrison SC. Structure of the dengue virus envelope protein after membrane fusion. Nature. 2004; 427(6972):313–9. https://doi.org/10.1038/nature02165 PMID: 14737159.

2. Kostyuchenko VA, Zhang Q, Tan JL, Ng TS, Lok SM. Immature and mature dengue serotype 1 virus structures provide insight into the maturation process. Journal of virology. 2013; 87(13):7700–7. https://doi.org/10.1128/JVI.00197-13 PMID: 23637416; PubMed Central PMCID: PMC3700294.

3. Kuhn RJ, Zhang W, Rossmann MG, Pletnev SV, Corver J, Lencehes E, et al. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell. 2002; 108(5):717–25. https://doi.org/10.1016/s0092-8674(02)00660-8 PMID: 11893341.

4. Zhang Y, Corver J, Chipman PR, Zhang W, Pletnev SV, Sedlak D, et al. Structures of immature flavivirus particles. The EMBO journal. 2003; 22(11):2604–13. https://doi.org/10.1093/emboj/cdg270 PMID: 12773377; PubMed Central PMCID: PMC156766.

5. Yu IM, Zhang W, Holdaway HA, Li L, Kostyuchenko VA, Chipman PR, et al. Structure of immature dengue virus at low pH primes proteolytic maturation. Science. 2008; 319(5871):1834–7. https://doi.org/10.1126/science.1153264 WOS:000254394000049. PMID: 18369148.

6. Kostyuchenko VA, Lim EX, Zhang S, Fibriansah G, Ng TS, Ooi JS, et al. Structure of the thermally stable Zika virus. Nature. 2016; 533(7603):425–8. https://doi.org/10.1038/nature17994 PMID: 27093288.

7. Sirohi D, Chen Z, Sun L, Klose T, Pierson TC, Rossmann MG, et al. The 3.8 Å resolution cryo-EM structure of Zika virus. Science. 2016; 352(6284):467–70. https://doi.org/10.1126/science.aaf5316 PMID: 27033547; PubMed Central PMCID: PMC4845753.

8. Dokland T, Walsh M, Mackenzie JM, Khromykh AA, Ee KH, Wang S. West Nile virus core protein; tetramer structure and ribbon formation. Structure. 2004; 12(7):1157–63. https://doi.org/10.1016/j.str.2004.04.024 PMID: 15242592.

9. Li T, Zhao Q, Yang X, Chen C, Yang K, Wu C, et al. Structural insight into the Zika virus capsid encapsulating the viral genome. Cell research. 2018; 28(4):497–9. https://doi.org/10.1038/s41422-018-0007-9 PMID: 29467384; PubMed Central PMCID: PMC5939043.

10. Ma L, Jones CT, Groesch TD, Kuhn RJ, Post CB. Solution structure of dengue virus capsid protein reveals another fold. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(10):3414–9. https://doi.org/10.1073/pnas.0305892101 PMID: 14993605; PubMed Central PMCID: PMC373476.

11. Shang Z, Song H, Shi Y, Qi J, Gao GF. Crystal Structure of the Capsid Protein from Zika Virus. Journal of molecular biology. 2018; 430(7):948–62. https://doi.org/10.1016/j.jmb.2018.02.006 PMID: 29454707.

12. Byk LA, Gamarnik AV. Properties and Functions of the Dengue Virus Capsid Protein. Annu Rev Virol. 2016; 3(1):263–81. https://doi.org/10.1146/annurev-virology-110615-042334 PMID: 27501261; PubMed Central PMCID: PMC4173333.

13. Prasad VM, Miller AS, Klose T, Sirohi D, Buda G, Jiann W, et al. Structure of the immature Zika virus capsid at 9 Å resolution. Nature structural & molecular biology. 2017; 24(2):184–6. https://doi.org/10.1038/nsmb.3352 PMID: 28067914; PubMed Central PMCID: PMC5296287.

14. Tan TY, Fibriansah G, Kostyuchenko VA, Ng TS, Lim XX, Zhang S, et al. Capsid protein structure in Zika virus reveals the flavivirus assembly process. Nature communications. 2020; 11(1):895. https://doi.org/10.1038/s41467-020-14647-9 PMID: 32060358; PubMed Central PMCID: PMC7021721.

15. Amberg SM, Nestorowicz A, McCourt DW, Rice CM. NS2B-3 protease-mediated processing in the yellow fever virus structural region: in vitro and in vivo studies. Journal of virology. 1994; 68(6):3794–802. PMID: 8189517; PubMed Central PMCID: PMC536864.

16. Ferlenghi I, Clarke M, Ruttan T, Allison SL, Schalich J, Heinz FX, et al. Molecular Organization of a Recombinant Subviral Particle from Tick-Borne Encephalitis Virus. Molecular Cell. 2001; 7(3):593–602. https://doi.org/10.1016/s1097-2765(01)00206-4 PMID: 11463384.

17. Martinez MG, Snapp EL, Perumal GS, Macaluso FP, Kielland M. Imaging the alphavirus exit pathway. Journal of virology. 2014; 88(12):6922–33. https://doi.org/10.1128/JVI.00592-14 PMID: 24696489; PubMed Central PMCID: PMC4504368.

18. Tellinghuisen TL, Perera R, Kuhn RJ. In vitro assembly of Sindbis virus core-like particles from cross-linked dimers of truncated and mutant capsid proteins. Journal of virology. 2001; 75(6):2810–7. https://doi.org/10.1128/JVI.75.6.2810-2817.2001 PMID: 11222705; PubMed Central PMCID: PMC115906.
19. Garoff H, Sjoberg M, Cheng RH. Budding of alphaviruses. Virus Res. 2004; 106(2):103–16. https://doi.org/10.1016/j.virusres.2004.08.008 PMID: 15567491.

20. Selzer L, Zlotnick A. Assembly and Release of Hepatitis B Virus. Cold Spring Harb Perspect Med. 2015; 5(12): a021394. https://doi.org/10.1101/cshperspect.a021394 PMID: 26527201; PubMed Central PMCID: PMC4665036.

21. de Haan CAM, Rottier PJM. Molecular Interactions in the Assembly of Coronaviruses. Virus Structure and Assembly. Advances in Virus Research. 2005; 64: 165–230. https://doi.org/10.1016/S0065-3527(05)64006-7 PMID: 16139595

22. Neuman BW, Adair BD, Yoshioka C, Quispe JD, Kuhn P, et al. Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. Journal of virology. 2006; 80(16):7918–28. https://doi.org/10.1128/JVI.00645-06 PMID: 16873249; PubMed Central PMCID: PMC1563836.

23. Klumpp K, Shimada T, Allweiss L, Volz T, Lutgehetmann M, Hartman G, et al. Efficacy of NVR 3–778, Alone and In Combination With Pegylated Interferon, vs Entecavir In uPA/SCID Mice With Humanized Livers and HBV Infection. Gastroenterology. 2018; 154(3):652–662.e8. Epub 24 Oct 2017. https://doi.org/10.1053/j.gastro.2017.10.017 PMID: 29079518.

24. Thenin-Houssier S, Valent ST. HIV-1 Capsid Inhibitors as Antiretroviral Agents. Curr HIV Res. 2016; 14(3):270–82. https://doi.org/10.2174/1570162x14999160224103555 PMID: 26957201; PubMed Central PMCID: PMC4785820.

25. Scaturro P, Trist IM, Paul D, Kumar A, Acosta EG, Byrd CM, et al. Characterization of the mode of action of a potent dengue virus capsid inhibitor. Journal of virology. 2014; 88(19):11540–55. https://doi.org/10.1128/JVI.01745-14 PMID: 25056895; PubMed Central PMCID: PMC4178822.

26. Byrd CM, Dai D, Grosenbach DW, Berhanu A, Jones KF, Cardwell KB, et al. A novel inhibitor of dengue virus replication that targets the capsid protein. Antimicrob Agents Chemother. 2013; 57(1):15–25. Epub 15 Oct 2012. https://doi.org/10.1128/AAC.01429-12 PMID: 23070172; PubMed Central PMCID: PMC3535982.

27. Martins IC, Gomes-Neto F, Faustino AF, Carvalho FA, Carneiro FA, Bozza PT, et al. The disordered N-terminal region of dengue virus capsid protein contains a lipid-droplet-binding motif. Biochem J. 2012; 444(3):405–15. https://doi.org/10.1042/BJ20112219 PMID: 22428600.

28. Paroutis P, Touret N, Grinstein S. The pH of the secretory pathway: measurement, determinants, and regulation. Physiology (Bethesda). 2004; 19:207–15. https://doi.org/10.1152/physiol.00005.2004 PMID: 15304635.