Computational characterization of viroporin structure and dynamics

Structural bases of proton dynamics in viroporins by multiscale modelling and simulation

***D. Rosenberger, F. Musil, C. Clementi**, Department of Physics, Freie Universität Berlin, Arnimallee 12, 14195 Berlin, Germany*

In Short

- We focus on the structure and dynamic of two small viral proteins: the M protein of the Dengue Virus (DENV) and of the West Nile Virus (WNV).
- These proteins adopt multimeric structures and form ion channels, and the blocking of their channel activity could potentially disrupt viral infection.
- We use a combination of structure prediction tools, coarse-grained simulation, and all-atom refinement incorporating experimental data to characterize the functional states and the conformational dynamics associated with the activity of these proteins.
- There are no vaccines nor cures against the DENG and WNV, and the characterization of the dynamics and function of protein M of these viruses presents an important opportunity as a potential drug target.

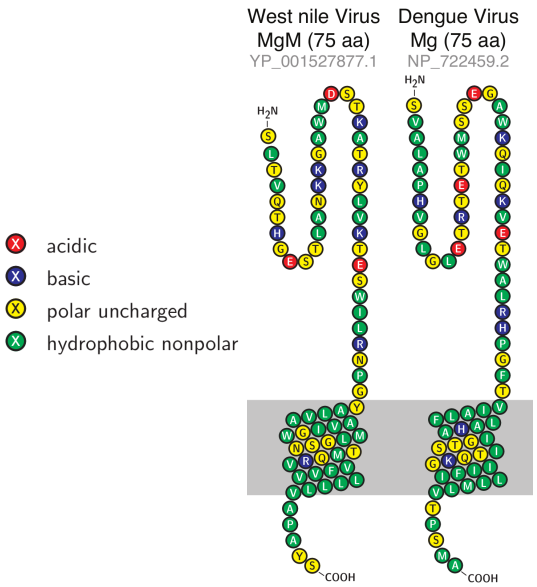

***Figure 1:** Sequence and predicted topology of the M protein of West Nile Virus and Dengue Virus. The shaded region indicates the presumed position of the lipid bilayer. The lengths of the proteins are stated in parentheses and the accession numbers are in grey. Figure adapted from 1*

Since the discovery of the M2 proton channel in the influenza A virus, and the ability to inhibit influenza by blocking this channel, a lot of attention has been devoted to the discovery and characterization of viroporins, small membrane proteins in viruses that are capable of ion channel activity.

The Dengue Virus (DENV) and the West Nile Virus (WNV) belong to the mosquito-borne cluster of the Flavivirus genus. The genome of Flaviviruses encodes three structural proteins – capsid, membrane (M), and envelope (E) – that constitute the virus particle. No biological function has yet been assigned to the M protein. The M proteins for DENV and WNV have strong similarities but also significant differences. In both DENV and WNV, the M protein has been recently shown to exhibit ion conductance 1, but the functional role of the M-mediated ion channel activity in these viruses is unknown.

The structure of protein M is not known for any Flavivirus, nor its oligomerization state in DENV and WNV. Based on sequence information, the M protein in both viruses has been predicted to have two trans-membrane helices (TMDs) and a stem helix when analyzed with three different transmembrane segment prediction programs; in addition, the protein's membrane incorporation was experimentally confirmed. The sequence and predicted topology are shown in Figures 1 and 2. The stem region (residues 22–37 in Dengue-2) is also helical and is partially buried in the outer lipid leaflet. In other studies, a single-TMD helix has been predicted for the M protein, implying that two membrane-associated forms may exist, potentially as a result of the tight turn ($\sim$3 aa) between the two TMDs (see Figure 2). A study of the cellular properties of protein M of the Dengue virus suggests that it forms oligomers when expressed as individual proteins in mammalian cells.

In the last few years, generative AI methods such as AlphaFold 2 have become popular tools for structure prediction. However, they are still not reliable in the prediction of multimeric complexes, nor do they provide dynamical information. We propose a methodology to characterize the structure and large structural rearrangements of the M protein of the DENV and WNV. For each system, we use AlphaFold to obtain an initial model for the protein monomer, then determine a stable multimeric assembly by means of a multiscale procedure. Different molecular dynamics (MD) simulations are performed with multiple monomer copies (ranging from 2 to 6) embedded in a lipid bilayer, with an available coarse-grained (CG) protein force field (MARTINI 3).

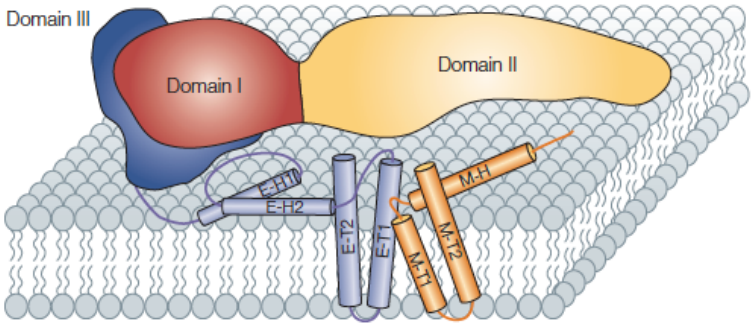

Figure 2: *Arrangement of the E and M proteins of Flavivirus in the lipid membrane. Domains I, II, and III of the E protein are colored red, yellow, and blue, respectively. The stem and transmembrane helices of the E (E-H1, E-H2, E-T1 and E-T2) and M (M-H, M-T1 and M-T2) proteins are shown in blue and orange, respectively. Figure adapted from 4.*

The use of a CG model allows us to simulate the timescale necessary to observe the assembly of the multimeric structures. Although the MARTINI model is not always quantitative accurate and it is not reliable for the simulation of globular proteins in solution, it is considered a safe choice for mechanistic studies of membrane proteins. We modify the MARTINI force field by using the AlphaFold confidence score in each residue position, as it has recently been shown to improve the performance and accuracy of the simulation 5. The CG simulations are then analyzed and metastable states are identified by using Markov State Models 6. Several (> 10) representative structures from each of the long-lived states are reconstructed in all-atom details and used as starting configurations for long (several μs) all-atom MD simulations to test the stability of the putative multimeric structures obtained with MARTINI.

With this multiscale approach, and using local computational resources, we have obtained exciting preliminary results strongly suggesting that the M protein of the Dengue virus forms a stable pore-forming trimer in lipid bilayers. We are now further testing these results by alternating rounds of simulation at the CG level for extensive sampling, and all-atom simulation for refinement. We expect this approach to yield a viable complex for the M protein of the Dengue virus, that can be tested experimentally by time-resolved pH-jump-initiated IR experiments on time scales from microseconds to seconds by our collaborators (J. Kozuch), and used for extensive all-atom MD and QM/MM simulations to study the ion permeability of the protein (in collaboration with H. Sun). We are also starting to repeat the same multiscale modeling framework for the M protein of WNV. The timescale involved, the size of the systems, and the numerous multimeric assemblies that are simulated at this stage of the project are only possible on high-performance computing clusters like the systems at the HLRN.

The complete characterization will be used to screen small molecules that can act as inhibitors for these viral channels for therapeutic application.

WWW

https://www.sfb1078.de/research/c7

More Information

[1] P. P. S. Tomar, R. Oren, M. Krugliak, I. T. Arkin. *Viruses*, **11**(7), 632, (2019). doi: 10.3390/v11070632

[2] Jumper, J., Evans, R., Pritzel, A. *et al. Nature* **596**, 583–589 (2021). doi:10.1038/s41586-021-03819-2

[3] L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, and S.-J. Marrink, *J. Chem. Theory Comput.*, **4**, 819-834 (2008). doi:10.1021/ct700324x

[4] Mukhopadhyay, S., Kuhn, R. and Rossmann, M. *Nat Rev Microbiol*, **3**, 13–22 (2005). doi: 10.1038/nrmicro1067

[5] A. Jussupow, and R. I. Kaila *J. Chem. Theory Comput.*, **19**(7) 1965–1975 (2023) doi: 10.1021/acs.jctc.2c01027

[6] F. Nüske, H. Wu, J.-H. Prinz, C. Wehmeyer, C. Clementi, F. Noé *J. Chem. Phys.*, **146**, 094104, (2017). doi:10.1063/1.4976518

Project Partners

Jacek Kozuch, Freie Universität Berlin; Han Sun, FMP Berlin

Funding

DFG SFB 1078 Project C07

DFG Subject Area

201-02 Biophysics