Biochemical and Biophysical Research Communications 425 (2012) 207–211

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

A molecular model and Monte Carlo simulation of flavivirus envelope building block

Dmitry I. Osolodkin [a,b,c], Liubov I. Kozlovskaya [b], Vladimir A. Palyulin [a,c,*], Vladimir M. Pentkovski [c], Galina G. Karganova [b], Nikolay S. Zefirov [a,c]

[a] Department of Chemistry, Lomonosov Moscow State University, Moscow 119991, Russia
[b] FSBI Chumakov Institute of Poliomyelitis and Viral Encephalitides RAMS, Moscow 142782, Russia
[c] iSCALARE Laboratory, Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia

ARTICLE INFO

Article history:
Received 3 July 2012
Available online 23 July 2012

Keywords:
Flavivirus
Implicit membrane
Molecular dynamics
Monte Carlo simulations
Envelope proteins

ABSTRACT

A molecular model of mature dengue virus envelope building block consisting of two E and two M protein subunits is constructed. Monte Carlo optimisation is performed for the model in the implicit membrane system. Interactions between the ectodomains and membrane parts of E and M proteins are studied, and possible role of conserved residues is suggested. The opening of E protein detergent binding site and its fixation in the open state by the M non-structured loop is observed.

1. Introduction

Such infections as dengue fever, West Nile fever, yellow fever, tick-borne encephalitis, etc., are a major health concern all over the world [1,2]. They are caused by viruses belonging to genus *Flavivirus* (family *Flaviviridae*), which are enveloped viruses, whose nucleocapsid is covered with a lipid membrane, which in turn is covered by the envelope protein E and membrane protein M, anchored in this membrane [3]. The outer surface of the mature virion is lined mainly by E protein ectodomain dimers forming the herringbone pattern in the mature fusion-ready state [4]. Viral fusion is initiated by low pH and leads to the formation of E ectodomain trimeric spikes [5], which attack the host membrane with their fusion loops (the flaviviral fusion process is thoroughly reviewed by Kaufmann and Rossmann [6]). Protein M is initially expressed as 166-residue prM protein, and during virion maturation in the host cells 91 'pr' residues are cleaved by cellular protease furin. These residues form a distinct domain which lies on the contact surface between fusion loop and domain III of neighbour E protein subunits, thus acting as a chaperone, preventing fusion loop from the premature detachment [7].

It has been shown for different flaviviruses that point mutations in the envelope proteins affect virus properties dramatically, sometimes making the virus low virulent upon peripheral inoculation in mice, unable to infect host cells or even not allowing it to assemble [8–11]. Structure-based mechanistic understanding of intrinsic machinery underlying these properties is required for rational design of new antiviral drugs, but such understanding is difficult due to the absence of high-resolution structures of the whole envelope or at least the whole building blocks forming this envelope. X-ray structures are available only for E protein ectodomain (sE) dimers [12,13] and trimers [5], whereas for E protein regions buried in the membrane, and for the M protein only low-resolution cryo-electron microscopy map is available [4].

With the help of this map and other X-ray structures we have constructed an atomic-scale model of a heterotetrameric building block of the mature fusion-ready high-pH dengue virus (DENV) serotype II envelope, consisting of two E subunits and two M subunits, and performed a Monte Carlo optimisation of this model in the implicit membrane followed by an analysis of the structural role of particular amino acid residues. This model will be used for the further studies of the flaviviral fusion.

2. Materials and methods

The templates used for the homology modelling of the heterotetramer are given in Table 1. Alignments were built with Clustal X 2.0.11 [14] and homology modelling was performed with Modeller 9.10 [15].

The first approximate model was built using 3C5X as the sE template. We selected the model where M protein disordered loop did not form knots with sE structure for thorough modelling using more general 1OAN template for sE. Additional restraints were applied to helicity of the well-established stem and anchor helices

* Corresponding author. Address: Department of Chemistry, Lomonosov Moscow State University, Moscow 119991, Russia. Fax: +7 495 939 02 90.
E-mail address: vap@qsar.chem.msu.ru (V.A. Palyulin).

0006-291X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2012.07.069

Table 1
Templates for the envelope proteins homology modelling.

PDB ID	Resolution, Å	Method	Role	Reference
1P58	9.50	Cryo-EM	General template for the whole sequence	[4]
1OAN	2.75	X-ray	High-resolution template for sE	[13]
3C5X	2.20	X-ray	Orientation of the M disordered loop	[29]

of the both proteins [4]. Stereochemical quality of the model was assessed by PROCHECK [16]; only 1.3% of all residues were in the generously allowed and disallowed regions of the Ramachandran plot, supporting the high quality of the model. The resulting model is shown in Fig. 1.

Monte Carlo (MC) simulation was performed in HIPPO, where a model for implicit solvent including the implicit membrane was implemented [17,18]. Trajectory analysis was performed with *analyse* from HIPPO package, VMD [19], and *cpptraj* [20]. The simulation protocol consisted of the following steps:

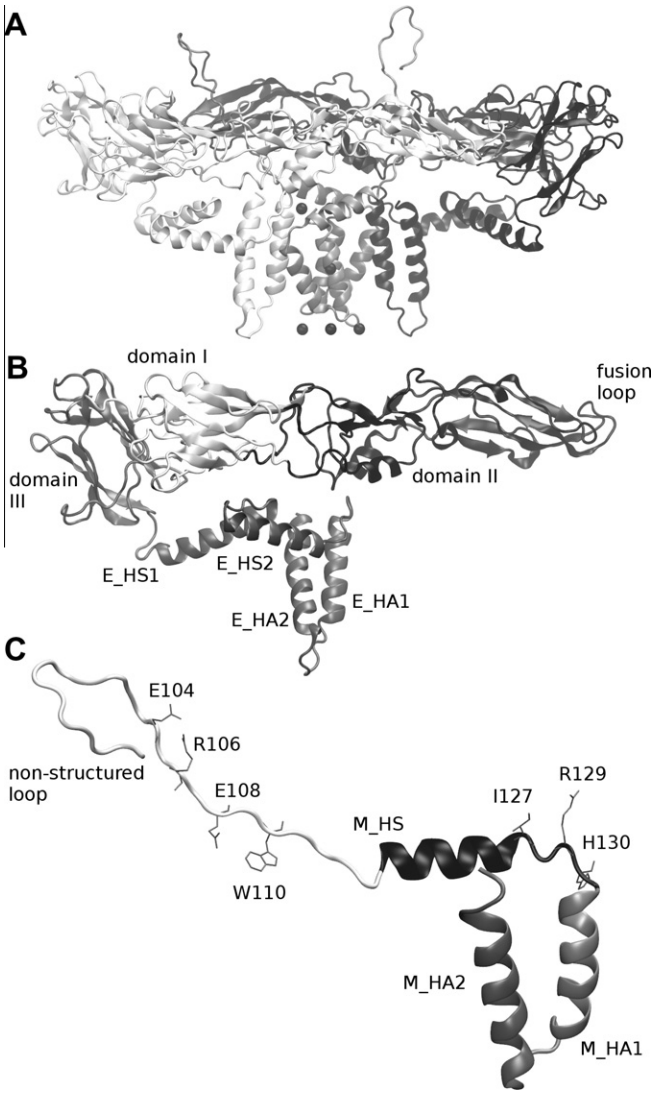

Fig. 1. (A) Overview of the starting model. Dark balls define the implicit membrane borders and middle plane. (B) Structural elements of the E protein. (C) Structural elements of the M protein. Stem helices are denoted HS, anchor helices are denoted HA.

1. Arbitrary assignment of the membrane central plane to the middle of anchor helices (membrane thickness was equal to 30 Å).
2. Translations/rotations (transrot) scan to find the optimal orientation of the model with respect to the membrane.
3. MC simulation for 2 M steps at 323 K.
4. Transrot scan for the resulting structure.
5. Production MC simulation: 4 runs for 2 M steps at 323 K, then 7 runs at 313 K to finish at 55 M steps after equilibration of the system.

System can be considered as equilibrated during the last 20 M steps, when change of the random seed between the runs leads to larger RMSD changes than can be observed during the run (Fig. 2A). Further simulation in the implicit solvent would be non-informative, since the influence of explicit water and lipid molecules would be much more prominent than the influence of internal protein features; small RMSD drift can be attributed to the elimination of errors introduced during the cryo-EM map fit and homology modelling. Nevertheless, certain conclusions can be made based on the results of this simulation.

3. Results and discussion

The initial orientation of the protein model with respect to the implicit membrane was chosen arbitrarily with two restraints: (i) ectodomain is parallel to the membrane plane and (ii) membrane central plane corresponds to the middle of the anchor helices. To get more realistic orientation of the model, a transrot scan was performed, and the lowest energy structure was analysed. In this structure stem helices are positioned along the membrane border, in accordance with experimental data [4,21], corresponding to their interaction with phosphate groups of the lipids, and the ectodomain is not perfectly parallel to the membrane (tilt angle 5°) since the template structure is not perfectly symmetric. Arginine residues at the ends of the E_HA1–E_HA2 loops (Fig. 1B) are extended and should interact with intrinsic membrane border via their guanidine groups. 2 M steps MC optimisation and repeated transrot scan do not affect the protein orientation significantly, and the resulting structure was used for further simulations.

The structure averaged over the last 8 M steps is shown in Fig. 3A. It closely resembles the starting structure, but certain differences can be observed. The main of them is movement of the E_HS2 stem helix toward the E protein ectodomain and formation of the contacts between them (Fig. 3B and C). In concordance with this, solvent-accessible surface area becomes smaller along the trajectory (Fig. 2B), and the distance between E_Leu425 and salt bridge E_Asp22–E_Lys284 decreases (Fig. 2C). Thus, the cleft between the stem and domain I shows the tendency to collapse. On the other hand, position of the M protein stem helix M_HS does not change substantially relatively to the E protein domain II.

Residue E_425 (426 in tick-borne viruses) is a host-specific determinant of flaviviruses [8]. It is located in a well-conserved sequence at the N-terminus of the loop between E_HS1 and E_HS2 [22] and interacts with the ectodomain via its side-chain. Nature of the residue in this position correlates with the species of the arthropod vector. Threonine residue in this position is characteristic for tick-borne encephalitis virus (TBEV) subtypes transmitted by *Ixodes persulcatus*. This residue is an obvious candidate to interact with the E_Asp22–E_Lys284 salt bridge; for example, it may disrupt this bridge or be clenched into a hydrogen-bound triad E_Asp22–E_Thr425–E_Lys284. Hydrophobic residues cannot form such interactions, thus leading to weaker interaction between the E protein stem loop and ectodomain.

The bending of ectodomain observed in previous TBEV sE molecular dynamics (MD) studies [9] is effectively absent

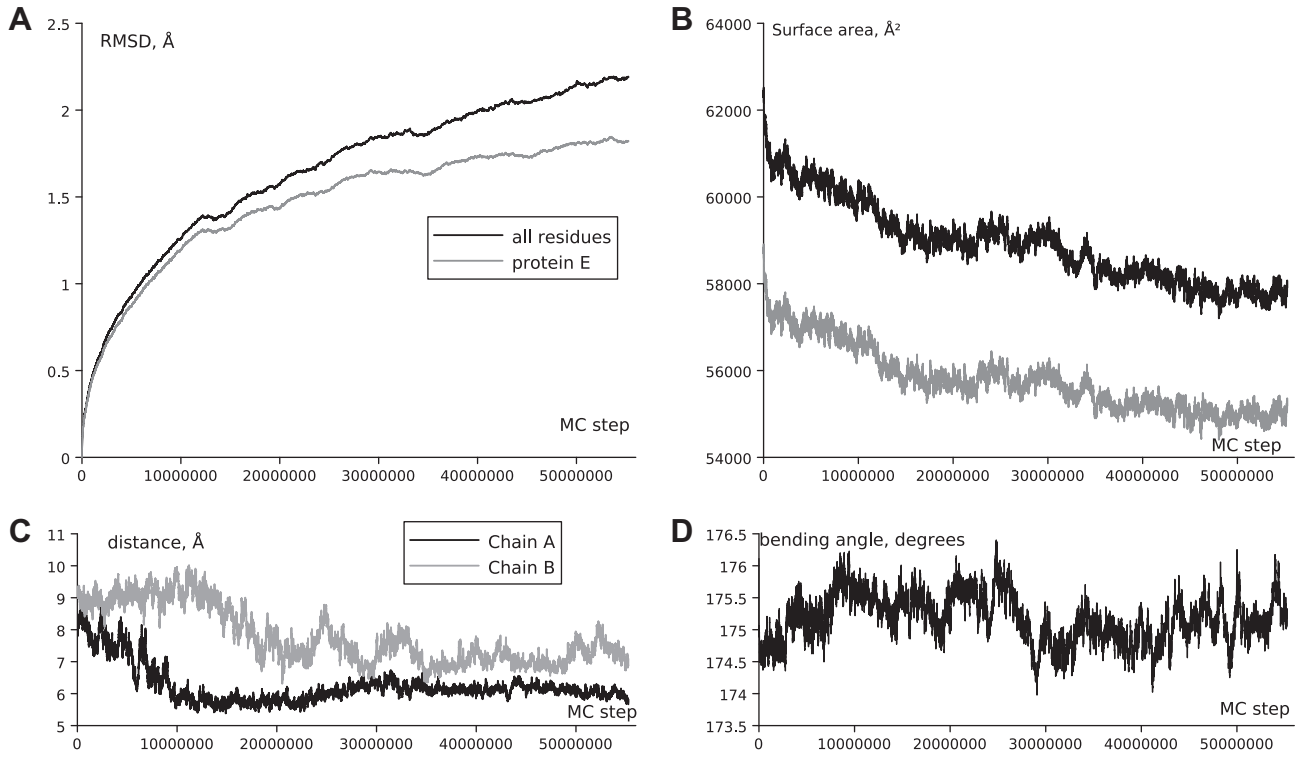

Fig. 2. Results of the trajectory analysis. (A) C_α atom RMSD for all residues (black) and E protein residues (grey). (B) Solvent-accessible surface area calculated by LCPO method [41] for all atoms (black) and with the M protein unstructured loop excluded (grey). (C) Distance between Leu425 and Asp22-Lys284 salt bridge (Chain A, black; Chain B, grey). (D) Bending angle for E protein ectodomain.

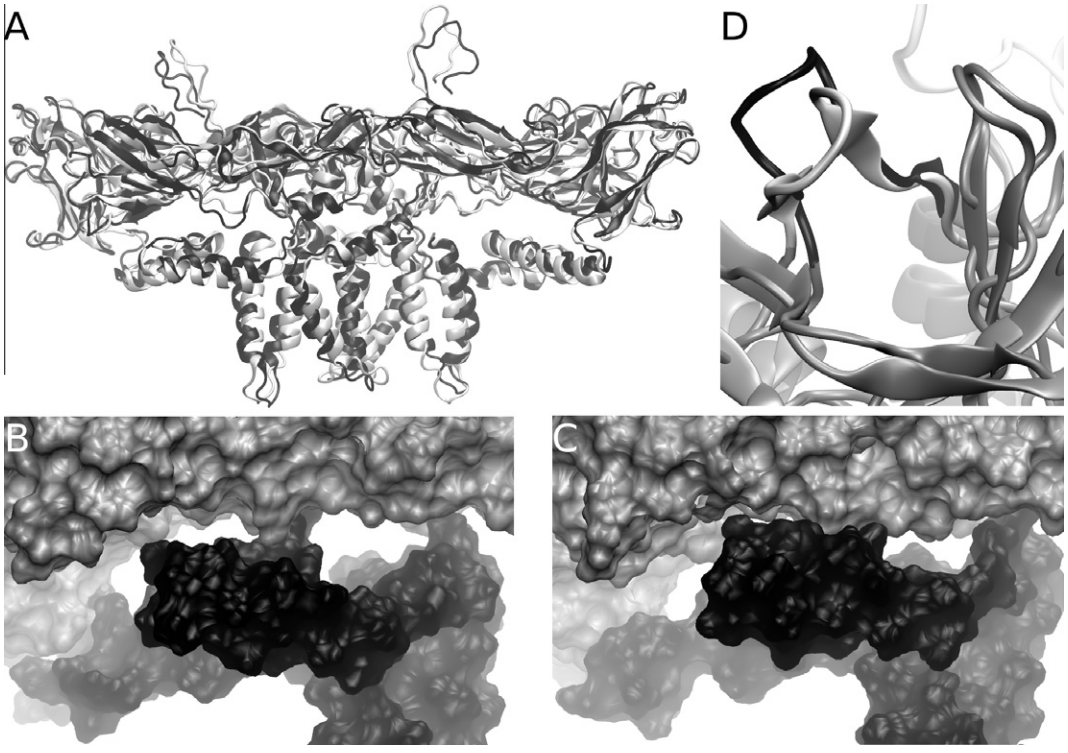

Fig. 3. (A) Spatial alignment of the initial model (white) and the model averaged over the last 8 M steps (black). (B) The cleft between E domain I (grey) and E_HS2 (black) in the initial model. (C) The contact between E domain I (grey) and E_HS2 (black) in the average model. (D) Open (black) and closed (light grey) conformations of the *kl* loop; the M protein is coloured white.

(Fig. 2D). The twist motion observed in our simulations of ectodomain dynamics as well as in independent ones [23] was also completely absent in the present simulation. Nevertheless, the helicity of E protein stem helix E_HS1 is not perfect along the

trajectory: this helix undergoes bending in the middle or slight disorder in the N-terminal part due to the pressure from the domain III and the presence of flexible E_Gly408 residue. In this simulation stem-anchor regions of E protein and M proteins are present, and their interaction with the ectodomains should be strong enough to avoid bending completely, making it a specific feature of soluble ectodomains.

A group of histidine residues (in particular, E_His144 (domain I), E_His317 (domain III) and E_His437 (stem)) is well conserved among flaviviruses and may participate in pH-dependent membrane fusion (so-called 'histidine switch' reviewed in [24]). A hydrogen bond is formed in our simulation between E_His437 and E_Gly190. The presence of a hydrogen bond donor/acceptor in E protein position 437 is important for fusion, whereas less similar substitutions affect the virion assembly process [25].

Residue E_His317 has been shown to play a determining role in the pH sensitivity of fusion [25] and has been suggested to stabilize E protein domain I/III interface in the low-pH conformation [26]. According to our simulation, such stabilization is achieved via hydrogen bonds formed by His sidechain NH as a donor and E_Ile6 and E_Arg9 backbone carbonyl oxygens as acceptors. The salt bridge E_Arg9–E_Glu373 is perfectly stable as opposed to earlier TBEV sE MD simulations [9], where it is significantly affected by E_His144, which forms a salt bridge with E_Glu373 in fusion trimers [24,26], or by E_His317. In our MC simulation, E_His144 formed a stable hydrogen bond with E_Asp42 sidechain but not with E_Thr353, which is located nearby. The absence of hydrogen bonding in the latter case may explain a destabilizing role of E_His144 residue in the domain I/III interface [26].

We have performed a first attempt to model the non-structured loop of the M protein. Folding of C-terminal residues 132–166 of M protein was hypothesised by Zhang et al. [4] according to cryo-EM map: one stem helix (M_HS) and two transmembrane (anchor) helices (M_HA1 and M_HA2) were suggested. This stem-anchor domain closely resembles the corresponding domain of E protein (E_HS2, E_HA1, and E_HA2), as shown in [27]. The structure of remaining residues 91–131 has never been studied properly; their invisibility in the cryo-EM map suggests non-structured character of this sequence. At the same time, N-terminus of this sequence must be accessible for the cellular protease when the premature virion is assembled [28]. The only way to build a model of this domain was to constrain its N-terminal residue position with the position of C-terminal residue of 'pr' domain available in the X-ray structure (PDB ID 3C5X [29]).

According to our model, the non-structured loop of the M protein goes through the 'holes' between the sE twin subunits, and the M backbone is located closer to the centre of the dimer. Interaction of sE with the M protein threaded through the hole leads to dramatic changes in the structure of E protein domain I/II interface. Namely, the *kl* loop, which closes the hydrophobic pocket accepting the detergent molecule in 1OKE structure [13], moves to make this pocket open (Fig. 3D; see also Supplementary video). In the sE template used in this study (PDB ID 1OAN), as well as in the other sE X-ray structures (1SVB [12]; 3P54 [30]), this pocket is usually closed, and its open conformation was obtained only in the complex with a detergent molecule. Since interaction with detergents prevents fusion, this pocket was successfully used in the search for new small-molecule fusion inhibitors by means of docking-based virtual screening [31–36]. The fixation of *kl* loop in open conformation by the M protein explains the availability of the pocket, because during the analysis of our earlier sE MD trajectories starting from closed-pocket structures ([9]; unpublished data) pocket opening was not observed. This finding also suggests that the closure of the *kl* loop may be an artifact of crystallisation conditions or innate feature of sE construct. The driving force of the loop opening is the penetration of conserved M_Arg106, whose sidechain is clenched between M_Glu104 and M_Glu108, into the space between E protein *kl* and *fg* loops. A salt bridge is formed between E_Lys202 and M_Glu104, and hydrophobic interaction appears between conserved M_Trp110 and E_Thr268.

M_His130 was shown to play an important role in DENV virion assembly, and its substitutions to polar residues are tolerated, but substitutions to non-polar residues disrupt normal functioning of the virus [10]. According to our simulation, in one of the simulated subunits this residue forms a hydrogen bond with E_Glu295, which also forms a salt bridge with M_Arg129. Positive charge in this position is conserved [28]. In the other subunit M_Arg129 forms a salt bridge with the C-terminal carboxyl of M_Thr166, and its sidechain may form hydrogen bond with conserved M_Glu124, which also forms a hydrogen bond with the sidechain of M_Gln121. These interaction patterns are important for virion formation and virus entry, as was shown by the alanine insertion between M_Arg129 and M_His130 [37].

The overall movement of M_HS is a slight submersion into the membrane relative to anchor helices. Disruption of the stem helix with proline residues [11] results in the impairment of contacts between the helix and E protein and thus impairs the virus entry. The stem helix is flanked by M_Ile127. Its interaction with E_Leu216 and hydrophobic methylene groups of E_Gln211 and E_Asp215 observed in our simulation can explain the requirement for a large hydrophobic sidechain in this position [11]. Proline substitutions in positions 120 and 123 of the M protein prevent the virus particle production [11]. According to our data, they do not form contacts with the E protein but form a hydrophobic face of the M protein stem helix that is directed to the E protein. Polar residues on the other side of the helix should interact with the charged lipid heads, and optimisation of these interactions leads to the submersion of the helix.

Despite the substantial similarity between flaviviruses, different residues may be important for different viruses. For example, our model does not allow us to explain the data on the importance of certain residues for functioning of Japanese encephalitis virus (JEV) [38–40] due to substantial differences between DENV and JEV envelope proteins [11,29,39].

In this work, we have constructed and optimised a model of flaviviral envelope building block, consisting of two E and two M protein subunits. Analysis of point mutation data supported the validity of the model. Of a special importance is the observation of the E protein *kl* loop fixation in the open state with the help of the M protein. Further analysis of the flavivirus envelope proteins' models will be performed with the help of explicit water and membrane simulations. Implicit solvent model, despite its rather simplistic nature, provided us with significant new information, and assessment of this information in a more rigorously constructed system is underway.

Acknowledgments

The authors are thankful to Daniel R. Roe for assistance with *cpptraj* calculations. This work was supported by the Grant from the Government of Russian Federation (Resolution No. 220 dated April 09, 2010).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2012.07.069.

References

[1] C.P. Simmons, J.J. Farrar, N. van Vinh Chau, B. Wills, Dengue, New Engl. J. Med. 366 (2012) 1423–1432.

[2] D. Růžek, G. Dobler, O.D. Mantke, Tick-borne encephalitis: pathogenesis and clinical implications, Trav. Med. Infect. Dis. 8 (2010) 223–232.
[3] R.J. Kuhn, W. Zhang, M.G. Rossmann, et al., Structure of dengue virus: implications for flavivirus organization, maturation, and fusion, Cell 108 (2002) 717–725.
[4] W. Zhang, P.R. Chipman, J. Corver, et al., Visualization of membrane protein domains by cryo-electron microscopy of dengue virus, Nat. Struct. Biol. 10 (2003) 907–912.
[5] S. Bressanelli, K. Stiasny, S.L. Allison, et al., Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation, EMBO J. 23 (2004) 728–738.
[6] B. Kaufmann, M.G. Rossmann, Molecular mechanisms involved in the early steps of flavivirus cell entry, Microb. Infect. 13 (2011) 1–9.
[7] I.-M. Yu, W. Zhang, H.A. Holdaway, et al., Structure of the immature dengue virus at low pH primes proteolytic maturation, Science 319 (2008) 1834–1837.
[8] L.I. Romanova, A.P. Gmyl, T.I. Dzhivanian, et al., Microevolution of tick-borne encephalitis virus in course of host alternation, Virology 362 (2007) 75–84.
[9] L.I. Kozlovskaya, D.I. Osolodkin, A.S. Shevtsova, et al., GAG-binding variants of tick-borne encephalitis virus, Virology 398 (2010) 262–272.
[10] M.J. Pryor, L. Azzola, P.J. Wright, A.D. Davidson, Histidine 39 in the dengue virus type 2 M protein has an important role in virus assembly, J. Gen. Virol. 85 (2004) 3627–3636.
[11] S.-C. Hsieh, G. Zou, W.-Y. Tsai, et al., The C-terminal helical domain of dengue virus precursor membrane protein is involved in virus assembly and entry, Virology 410 (2011) 170–180.
[12] F.A. Rey, F.X. Heinz, C. Mandl, C. Kunz, S.C. Harrison, The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution, Nature 375 (1995) 291–298.
[13] Y. Modis, S. Ogata, D. Clements, S.C. Harrison, A ligand-binding pocket in the dengue virus envelope glycoprotein, Proc. Natl. Acad. Sci. USA 100 (2003) 6986–6991.
[14] M.A. Larkin, G. Blackshields, N.P. Brown, et al., Clustal W and Clustal X version 2.0, Bioinformatics 23 (2007) 2947–2948.
[15] A. Šali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[16] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK – a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[17] J.P. Ulmschneider, M.B. Ulmschneider, Hippo 1.0, Biowerkzeug Limited (2008).
[18] J.P. Ulmschneider, M.B. Ulmschneider, Sampling efficiency in explicit and implicit membrane environments studied by peptide folding simulations, Proteins 75 (2009) 586–597.
[19] W. Humphrey, A. Dalke, K. Schulten, VMD – visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[20] D.A. Case, T.E. Cheatham, T. Darden, et al., The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.
[21] H. Nemésio, F. Palomares-Jerez, J. Villalaín, The membrane-active regions of the dengue virus proteins C and E, Biochim. Biophys. Acta 1808 (2011) 2390–2402.
[22] K. Stiasny, S.L. Allison, A. Marchler-Bauer, C. Kunz, F.X. Heinz, Structural requirements for low-pH-induced rearrangements in the envelope glycoprotein of tick-borne encephalitis virus, J. Virol. 70 (1996) 8142–8147.
[23] D.S. Mueller, T. Kampmann, R. Yennamalli, et al., Histidine protonation and the activation of viral fusion proteins, Biochem. Soc. Trans. 36 (2008) 43–45.
[24] K. Stiasny, R. Fritz, K. Pangerl, F.X. Heinz, Molecular mechanisms of flavivirus membrane fusion, Amino Acids 41 (2011) 1159–1163.
[25] R. Fritz, K. Stiasny, F.X. Heinz, Identification of specific histidines as pH sensors in flavivirus membrane fusion, J. Cell Biol. 183 (2008) 353–361.
[26] M.K. Prakash, A. Barducci, M. Parrinello, Probing the mechanism of pH-induced large-scale conformational changes in dengue virus envelope protein using atomistic simulations, Biophys. J. 99 (2010) 588–594.
[27] M.V. Bogachek, E.V. Protopopova, V.A. Ternovoi, et al., Immunochemical properties of the West Nile virus prM protein and the C-terminal fragment of the M protein, Mol. Biol. 41 (2007) 5–13.
[28] M.R. Holbrook, H. Wang, A.D.T. Barrett, Langat virus M protein is structurally homologous to prM, J. Virol. 75 (2001) 3999–4001.
[29] L. Li, S.-M. Lok, I.-M. Yu, et al., The flavivirus precursor membrane-envelope protein complex: structure and maturation, Science 319 (2008) 1830–1834.
[30] V.C. Luca, J. AbiMansour, C.A. Nelson, D.H. Fremont, Crystal structure of the Japanese encephalitis virus envelope protein, J. Virol. 86 (2012) 2337–2346.
[31] J.-M. Yang, Y.-F. Chen, Y.-Y. Tu, K.-R. Yen, Y.-L. Yang, Combinatorial computational approaches to identify tetracycline derivatives as flavivirus inhibitors, PLoS ONE 2 (2007) e428.
[32] Z. Zhou, M. Khaliq, J.-E. Suk, et al., Antiviral compounds discovered by virtual screening of small-molecule libraries against dengue virus E protein, ACS Chem. Biol. 3 (2008) 765–775.
[33] Z. Li, M. Khaliq, Z. Zhou, et al., Design, synthesis, and biological evaluation of antiviral agents targeting flavivirus envelope proteins, J. Med. Chem. 51 (2008) 4660–4671.
[34] T. Kampmann, R. Yennamalli, P. Campbell, et al., *In silico* screening of small molecule libraries using the dengue virus envelope E protein has identified compounds with antiviral activity against multiple flaviviruses, Antiviral Res. 84 (2009) 234–241.
[35] M.K. Poh, A. Yip, S. Zhang, et al., A small molecule fusion inhibitor of dengue virus, Antiviral Res. 84 (2009) 260–266.
[36] A.S. Mayhoub, M. Khaliq, C. Botting, et al., An investigation of phenylthiazole antiflaviviral agents, Bioorg. Med. Chem. 19 (2011) 3845–3854.
[37] A. Op De Beeck, R. Molenkamp, M. Caron, et al., Role of the transmembrane domains of prM and E proteins in the formation of yellow fever virus envelope, J. Virol. 77 (2003) 813–820.
[38] Y.-J. Lin, S.-C. Wu, Histidine at residue 99 and the transmembrane region of the precursor membrane prM protein are important for the prM-E heterodimeric complex formation of Japanese encephalitis virus, J. Virol. 79 (2005) 8535–8544.
[39] Y.-J. Lin, J.-G. Peng, S.-C. Wu, Characterization of the GXXXG motif in the first transmembrane segment of Japanese encephalitis virus precursor membrane (prM) protein, J. Biomed. Sci. 17 (2010) 39.
[40] C.C. Maier, S. Delagrave, Z. Zhang, et al., A single M protein mutation affects the acid inactivation threshold and growth kinetics of a chimeric flavivirus, Virology 362 (2007) 468–474.
[41] J. Weiser, P.S. Shenkin, W.C. Still, Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO), J. Comput. Chem. 20 (1999) 217–230.