The effects of urea and of pH on protein structure studies by molecular dynamics simulation
Mueller, Daniela Sung-Mi

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
6.1 Introduction

The envelope protein E of a flavivirus is composed of a soluble ectodomain and a C-terminal, membrane-associated domain, the so-called stem-anchor region (Fig. 5.3). The ectodomain is a large structure of approximately 400 residues, suspended on the surface of the viral membrane (Fig. 3.2A). The truncated and solubilised fragment sE contains only the ectodomain. The stem-anchor region, or membrane domain, consists of approximately 100 residues. Its secondary structure was predicted to contain two membrane-peripheral helices, the stem helices H1 and H2, and two trans-membrane helices, the anchor helices T1 and T2, that are connected by flexible linkers [Zhang et al. 2003]. The flaviviral envelope proteins are structurally and functionally similar to the envelope proteins of the alphaviruses. The two groups are therefore collectively referred to as class II viral fusion proteins.

At low pH the E protein dimers dissociate into monomers and assemble into trimers [Allison et al. 1995; Modis et al. 2004]. The E trimer and the sE trimer both are more stable than the corresponding dimers. Furthermore the E trimer is more temperature-stable than the sE trimer [Stiasny et al. 2005]. To date the exact mechanism by which flaviviral E proteins induce fusion is not known, nor the atomic structure of the stem-anchor region. Therefore it remains to be shown what the specific function of the stem-anchor region is, and whether it plays an essential role in the fusion mechanism, other than attaching the E protein to the viral membrane.

The existing models of class II viral fusion assume that the active form of E is trimeric
6.1 Introduction

[Harrison 2008; Jardetzky & Lamb 2004], in analogy to the active, trimeric forms of alphaviral and influenza viral fusion proteins. However, there is no experimental evidence that in flaviviruses trimerisation occurs before fusion, and therefore it is unclear whether trimerisation is a requirement for the activation of flaviviral fusion proteins. In fusion assays of tick-borne encephalitis virus (TBEV), dengue virus (DEN2) and their respective recombinant subviral particles, pre-exposure to low pH resulted in the loss of fusion activity [Corver et al. 2000; Stiasny et al. 2002]. Due to the rapid and irreversible inactivation of flaviviral E protein after exposure to low pH, the post-fusion crystal structures of flaviviral sE proteins (PDB-IDs: 1OK8 [Modis et al. 2004], 1URZ [Bressanelli et al. 2004]) most likely represent the inactive form of the E protein ectodomain.

The kinetics of fusion were found to differ significantly between influenza virus, the alphavirus Semliki Forest virus (SFV), and the flavivirus TBEV. From this it was concluded that the fusion active state of the TBE viral E protein may fundamentally differ from the active states of the envelope proteins in influenza and Semliki Forest virus [Corver et al. 2000]. Furthermore, in TBEV the kinetics of inactivation are slow compared to the kinetics of fusion, which suggests that the inactivation may involve additional conformational changes in the E protein [Corver et al. 2000]. The trimerisation step would be a good candidate for a slow, concluding step in the conformational change of the E protein. Thus the notion that inactivation results from trimerisation fits the kinetics of the E protein.

The inactivity of the trimer might be due to: i) the inability of the fusion peptide to insert into the target membrane, due to crowding among the three fusion loops, resulting in restricted mobility, unfavourable positioning with respect to the lipid head groups, and repulsion from the membrane; ii) the simple fact that the trimer is held upright on the viral surface by symmetry constraints and neighbouring trimers, which act as spacers between the viral and target membranes and thereby prevent fusion. Point ii is in line with a proposal by Corver et al. [2000], that “the fast rate of fusion and the low activation energy of this [fusion] process [...] may be a consequence of the flat orientation of the E protein [on the viral surface], which, by lying parallel to the interacting membranes, would facilitate the establishment of direct molecular contact between them.” They concluded that the trimer formed under low-pH conditions may represent a final fusion-inactive structure, and that a minor conformational change may be sufficient for fusion [Corver et al. 2000]. Based on these considerations, this chapter investigates the minimal structural requirements for the conformational change of flaviviral E protein.

The exact structural dynamics of flaviviral E protein and any associated effects on

---

1 Although influenza viral envelope proteins were classified as class I viral fusion proteins, they have similar requirements for fusion and are also activated by low pH [White et al. 2008].

2 The X-ray crystal structures available of flaviviral E proteins in the post-fusion conformation were obtained from crystals of the soluble ectodomain sE.
the viral or the host membrane are unknown. However, some general functions of fusion-mediating proteins are known and include: the local bending of the membranes in order to minimise inter-membrane repulsion; to induce local dehydration at the membrane contact; to promote stalk formation by generating bilayer stress [Chernomordik & Kozlov 2003]. In fact, simulation studies of membranes [Leontiadou et al. 2007; Marrink & Mark 2004] and of bilayer fusion [Noguchi & Takasu 2001] support the so-called stalk-pore hypothesis of membrane fusion [Zimmerberg & Chernomordik 1999], in which during the initial stage of fusion the contacting leaflets of the fusing membranes connect and form a stalk.

The propensity for two membranes to fuse depends on the spontaneous curvature of the membranes, which in turn depends on the lipid composition [Chernomordik & Kozlov 2003]. In experiments with liposomes the absence of specific lipids impaired viral fusion [Lee et al. 2008; Stiasny & Heinz 2004; Stiasny et al. 2003]. Experimental [Fujii et al. 1993; Murata et al. 1991] and simulation studies [Pécheur et al. 1999] have shown that the interaction with small amphiphilic peptides can alter the spontaneous curvature of a membrane and lead to fusion. Larger proteins also induce curvature in membranes, for example the so-called
BAR domains (bin, amphiphysin and Rvs, members of the amphiphysin protein family) [Peter et al. 2004; Saarikangas et al. 2009], which bind preferentially to curved membranes and are found in proteins that regulate the remodeling of membranes, e.g. the shaping of compartments and organelles, or the recycling of synaptic vesicles [Gallop & McMahon 2005; Peter et al. 2004; Saarikangas et al. 2009]. MD simulations predicted that the N-BAR domain of amphiphysin induces a mean curvature of $0.15 \text{ nm}^{-1}$ [Ayton et al. 2007].

In MD simulations of the DEN2 sE dimer double-protonation of all the histidine residues led to a conformational change (see Chapter 5), while the protein remained dimeric. Specifically, hinge motions between the domains resulted in domain I to project towards where the membrane would be located in situ (Fig. 6.1) and predict interactions between the protein and the viral membrane. In the current chapter new models of intermediate conformations of the E protein are introduced, some of which include the membrane domain. The conformational change observed during the MD simulation of the protonated sE dimer is evaluated in terms of interactions between the ectodomain and the stem-anchor region. The implications of these interactions for the activation of flaviviral E proteins are discussed. A fusion model is proposed, in which fusion occurs before trimerisation. This fusion model is supplemented by a structural model of an intermediate after acidification, that was generated by MD simulation.

### 6.2 Methods

**Sequence variability analysis** The sequence variability of the full-length E protein was determined for the 28 flaviviral E protein sequences of the conserved domain (CD) superfamily pfam00869 that also code for the membrane domain of E [Marchler-Bauer et al. 2009]. The sequence variability was determined as the Shannon entropy [PVS; Shannon 1948]

$$H(l) = -\sum_{r=1}^{20} f(r,l) \log_2 f(r,l)$$  \hspace{1cm} (6.1)

The sequence similarity diagram was generated using the WebLogo web server [Crooks 2009; Crooks et al. 2004; Schneider & Stephens 1990].

The relative frequency $f$ of a residue type $r$ at position $l$ and the measure for sequence information $R_{seq}$ determine the height $L$ of the letter of the residue type $r$ in the WebLogo diagram

$$L(l) = f(r,l) R_{seq}(l)$$  \hspace{1cm} (6.2)

and is given in [bits per residue] [Schneider & Stephens 1990]. $R_{seq}$ is derived from the corre-
sponding Shannon entropy $H$ and a correction term $e(n)$ for small sample sizes $n$ [Schneider et al. 1986]

$$R_{seq}(l) = E(H_{nr}) - H(l) = \log_2 20 - (H(l) + e(n))$$

$(n = 28 \Rightarrow e(n) \approx 0.5)$. $H_{nr}$ is the uncertainty of obtaining a particular combination of $n$ residues of types $r$

$$H_{nr} = -\sum_{r=1}^{20} \left( \frac{n_r}{n} \right) \log_2 \left( \frac{n_r}{n} \right)$$

where $n = \sum_{r} n_r$ is the number of sequences without gaps, and $P_r$ is the general frequency of residue type $r$ [Schneider et al. 1986], which is simplified to equiprobability for all residue types, i.e. $P_r = \frac{1}{20} \forall r$ [Schneider & Stephens 1990]. The average of the uncertainty $H_{nr}$ is given by the expectation value

$$E(H_{nr}) = \sum_{all\ nr} P_{nr} H_{nr}$$

weighted by the probability $P_{nr}$ of obtaining the effective combination:

$$P_{nr} = \frac{n!}{\prod_{r} n_r!} \prod_{r} P_r^{n_r}$$

**Atomistic model of the full-length E protein** An atomistic, full-residue model of the membrane domain was generated from the C$\alpha$-trace of a cryo-EM structure (PDB-ID: 1P58 [Zhang et al. 2003]) using the Geno3D web server [Combet et al. 2002]. Atomistic structures of the ectodomain of the E dimer were obtained from X-ray crystal structures (PDB-IDs: 1OAN and 1OKE [Modis et al. 2003]) and from the configuration of an MD simulation (Chapter chapter 5). These were combined with the model of the membrane domain, by superposition onto the cryo-EM structure of the full-length E dimer, 1P58.
6.3 Results

The low-pH-dependent conformational change of the E protein

The pre-fusion and the post-fusion conformation of the DEN2 E protein ectodomain were compared to determine the differences in the tertiary structure that result from the conformational change. In order to visualise the rearrangement of the domains, the crystal structures of the backbone in the two conformations of sE were superimposed (PDB-IDs: 1OAN [Modis et al. 2003], 1OK8 [Modis et al. 2004]); the two superpositions are shown in Figure 6.2A.

In the left panel of the figure the superposition of domain I, the so-called organising domain of E, shows that domains II and III both move in similar directions relative to domain I. Accordingly, in the superposition of domains II and III, shown in the right panel, the most obvious difference is the position of domain I. In the following these two superpositions are called “SI” and “SII+III”, after the respective domains used for the fits.

In the superposition SII+III the main difference between the two conformations is the position of domain I, which appears to be released or ejected from between domains II and III during the conformational change. Figure 6.2B shows a model of a possible intermediate conformation (orange) after the release of domain I. This model is a hybrid of the two superimposed structures, assembled from domain I of the post-fusion structure and domains II and III of the pre-fusion structure (orange). In the figure the pre-fusion structure (yellow), the post-fusion structure (red) and the hybrid model (orange) are compared in various dimeric assemblies. These were assembled by superposition onto the subunits of the pre-fusion dimer (Fig. 3.2A p. 33).

Another noticeable difference between the pre-fusion and the post-fusion structure in SII+III (Fig. 6.2A) is the position of domain II (yellow). Domain II is shifted by approximately 9 nm along the long principal axis of the dimer towards domain I (red). Based on this finding, the hybrid model was modified in a second modelling step (Fig. 6.3), where domain II (orange) was shifted by 9.0 nm (green), towards domain I (red). As can be seen in the figure, this shift eliminates the steric clash between the fusion loop (blue) and domain III (orange), that is still present in the opposite homomer where no shift was applied. In addition, the shift separates the helices in the central region of the subunit interface.

In order to evaluate the fits of the superpositions (Fig. 6.2A), the RMSD between the superimposed structures was calculated. The RMSD of all the backbone atoms was 2.1 nm in SI and 1.3 nm in SII+III. The shift of the center of mass of domain III between the superimposed structures was 3.4 nm in SI and in SII+III it was 1.2 nm, i.e. less than half of the shift in SI. The shift of the C-terminus was 4.0 nm in SI and 2.1 nm in SII+III; the C-terminus is located on the surface of domain III. In terms of these three criteria, SII+III matched the
conformations more closely than SI and therefore seems the better choice for modelling the transition between the two conformations.

**Interaction between the ectodomain and the stem-anchor region**

An atomistic model of the full-length E dimer was obtained by superposition of the pre-fusion crystal structure of the E protein dimer and of an atomistic model of the membrane domain onto the cryo-EM structure of the mature E protein. The resulting structure is shown in Figures 6.1A and 6.4. In the cryo-EM structure residues 395–399, which link the C-
terminus of the ectodomain to the C-terminal membrane domain, were not resolved. The RMSD between the Cα-atoms of 1OAN and 1P58 was 0.5 nm. The conservation of the full-length E protein was analysed to predict conserved interactions at the interface between the ectodomain and the stem-anchor region. The conserved domain database CDD identified 38 representative sequences in the flaviviral glycoprotein superfamily (pfam00869 [Marchler-Bauer et al. 2009; Marchler-Bauer & Bryant 2004]). Twenty-eight of these sequences also code for the stem anchor region and were analysed in terms of sequence variability; the result is presented as a WebLogo diagram in Figure 6.5.

For the prediction of salt bridges between the ectodomain and the membrane domain and of hydrophobic interactions between the protein and the viral membrane, the conservation of charged, or large and hydrophobic residues at the protein surface was analysed. Figure 6.6 indicates the degree of conservation of the residues at the surface of the E protein: Colours indicate full conservation, lighter shades of grey higher conservation, and darker shades less conservation. The surface of the ectodomain facing the viral membrane (Fig. 6.6C) was generally more conserved than the external surface. Figure 6.6C shows the membrane domain rotated by 180° around a vertical axis into an open-book presentation alongside the ectodomain. On the surface of the ectodomain, Asp10, Phe11, Glu26, Lys110, His244, His282 and His317 were fully conserved. These residues cluster around the interfaces between domains I, II’ and III, opposite the stem helices (Fig. 6.4). Trp212, Trp220 and Phe240, which are located on the surface at the hinge region between the base and the elongated part of domain II (Fig. 6.2, yellow and green respectively) were fully conserved. In the stem helices and the CS linker, Arg407, Arg411, Asp417, Trp420, Asp421, Phe422, Lys434, His437 and Phe448 were conserved. In the atomistic model of the stem-anchor region, Asp421 and His437 formed a salt bridge between the CS linker and the stem helix H2.
The CS linker is highly conserved [Zhang et al. 2003] (Fig. 6.5) and contains a sequence of small residues 423–427, GSLGG.

In Figure 6.6C it can be seen that the stem helix H1 and the CS linker align with highly conserved residues on the surfaces of domains I and III. The conserved, charged residues Arg411 and Asp417 of H1 align with the conserved, complementary charged residues Asp10, Glu26 and His282 on domain I (Fig. 6.6). In Figure 6.4 it can furthermore be seen that these conserved residues are located along the interfaces of domains I (rose) and III (blue) with domain II’ (orange), the fusion domain of the opposite subunit. On the surface of domain I (rose) a long groove extends around approximately one third of the circumference of the domain; in Figure 6.4 red boxes highlight the groove. The groove begins approximately opposite the C-terminus of stem helix H1 and leads along a series of highly conserved hydrophobic residues, Val24, Cys185 and Cys285, across the β-sheet BIH.

**Structural intermediate in the conformational change of the full-length E protein**

An MD simulation was performed to test the effect of double-protonation of all the histidine residues on the structure of the ectodomain. The stem-anchor region was modelled into the configuration of the sE dimer after 70 ns of simulation to determine whether the conformational changes observed were relevant to interactions with the membrane domain of E. The model was generated by superposition onto the cryo-EM structure of domains II and III of the dimer, in line with the model SII+III established above (p. 108). The model is shown in

---

**Figure 6.4** The membrane-facing surface of the full-length E protein: the ectodomain (surface) and the Cα-trace of the membrane domain (blue tube). From left to right: The structure is rotated in steps of 30 ° around a vertical axis, from the view of the surface facing the membrane to the side view. Red boxes highlight the extensive groove on the surface of domain I. The grey-scale colour gradient indicates the conservation of the residues (cf. Fig. 6.6), white corresponds to full conservation. Domain colouring scheme: rose – domain dl, yellow/orange – dll of the two respective subunits, blue – dlll.
Figure 6.5 Sequence similarity among 28 flaviviral E proteins. The height of a letter at a particular position is proportional to the frequency of the corresponding residue type among the aligned sequences and to the sequence information in the alignment and is given in bits [Crooks 2009; Crooks et al. 2004; Schneider & Stephens 1990]. The residue numbering is based on the dengue viral type 2 E protein, gi|31615787. The letters are coloured according to the residue type: red – acidic, blue – basic, orange – histidine, green – polar, rose – hydrophobic, black – large and hydrophobic.
Figures 6.1B–C, in which can be seen that in both subunits the domains I arch out towards the respective membrane domain. Furthermore, structural elements of domains I and II extend between the stem helices and the anchor helices respectively.

The steric overlaps in the superposition, between domain I and the membrane domain, are due to the deformation of the protein and the bias of the superposition by the reference domains dII and dIII. In order to alleviate these overlaps, the positions of the membrane domains were modelled individually for each subunit, as shown in Figure 6.7. The internal structures of the stem-anchor region were preserved. In both subunits of this model, domain I interacts with the CS linker and separates the stem helices. Due to the amphipathic nature of the stem helices [Zhang et al. 2003], they are expected to remain suspended in the head group region of the membrane. The stem helix H1 would be constrained by the C-terminus of the ectodomain (purple beads), while the stem helix H2 may move at the end of the flexible CS linker.
Figure 6.7 previous page Model of an intermediate from the MD simulation of the E protein after histidine protonation, interacting with the membrane domain. The configuration of the ectodomain sE (surface) after 70 ns of MD simulation was superimposed onto domains II and III of the cryo-EM structure of the E dimer. A) Side view. B) Cross-eyed stereo view of that side of the E protein that faces the membrane. In B the rigid structures of the two stem-anchor regions (blue tubes) of the dimer were modelled individually to fit the configuration of the respective subunit of the ectodomain. Fully conserved residues are represented by beads coloured according to residue type: red – acidic, blue – basic, green – polar, white – hydrophobic; C-terminus – purple. Domain colouring scheme: rose – domain dl, yellow/orange – dII, blue – dIII.
6.4 Discussion

The low-pH-dependent conformational change of the E protein

Until now the conformational change from the pre-fusion to the post-fusion structure has commonly been described as a motion of domain III relative to the other two domains (Figs. 3.1, 6.2A left) [Harrison 2008; Modis et al. 2004]. While this is true, one must consider that the process occurs in the envelope environment and possibly involves the viral membrane. Here, an alternative view on the conformational change is introduced (Figs. 6.2, 6.3), that considers the restrictive membrane environment and other experimental findings. The new model of the conformational change presented here is based on the same structures as the previous model, but uses a different set of atoms as reference for the domain motions. This condenses the conformational change to a motion of domain I, which is released from its meta-stable position between domains II and III (SII+III, Fig. 6.2A right).

From structural comparisons similar to the superposition SI (Fig. 6.2A left) it was previously concluded that the conformational change required more lateral space than is available in the packed envelope. In search of a solution to this problem it was even proposed that a radial expansion of the viral membrane might accommodate the conformational change at low pH [Bressanelli et al. 2004], but this seems unlikely from a physical and chemical perspective. In contrast the superposition onto domains II and III, SII+III (Fig. 6.2A right), shows that the post-fusion structure fits into the initial flat arrangement of the pre-fusion dimer, where it would occupy a similar area on the viral membrane as the pre-fusion structure (Fig. 6.2B). Therefore no additional space is needed in this model of the conformational change of the subunit, for the release of domain I.

The structural rearrangements in the new model are less dramatic than in the previous model and may therefore occur on a faster time scale. This stands in agreement with the extraordinarily fast rate of fusion from which Corver et al. [2000] inferred that the TBE viral E protein was active in a flat orientation and that the activation may therefore require only a minimal conformational change. The new model presented here reconciles a flat orientation with the complete conformational change concluded from the crystal structures (excluding the oligomeric rearrangement into trimers; Figs. 6.2, 6.3). Even though the protein may not remain in a perfectly flat orientation during the entire fusion process, the new model shows that the domain rearrangements could, in theory, be completed to a large degree in the flat dimer.

The new model furthermore suggests that the dissociation of the dimer and the spiking of the envelope might occur after most of the domain rearrangements have already taken place. This matches the kinetics of the E protein, which suggest that after the fast activation
a slow conformational change follows that leads to inactivation [Corver et al. 2000]. The assembly of the trimer is likely to be a time-consuming process and may therefore correspond to the slow inactivation step. Fusion would occur in between the fast and the slow step, after an initial conformational change into a yet unknown intermediate state. A model for the active intermediate conformation of flaviviral E protein, based on the results of MD simulations, (Chapter chapter 5) is proposed in the following sections.

**Interaction between the ectodomain and the stem-anchor region**

In the superposition of the intermediate configuration of sE, obtained from the simulation of sE after histidine protonation, onto the structure of the full-length E protein, the juxtaposition of the deformation of domain I with the membrane domain of the cryo-EM structure is striking (Figs. 6.1B–C, 6.7). The strip of highly conserved residues on domains I and III is directly opposite the stem helix H1 in the pre-fusion conformation and strongly suggests that this is an important interface (Fig. 6.4). The alignment of H1 with the interface between domains I and III and the other subunit suggests that interactions between H1 and domains I and III may play a role in the dissociation of the fusion domain.

The interactions with the stem helices H1 and H2 that were predicted from the intermediate configuration suggest that the deformation of domain I observed during the simulation might affect the positions of the helices (Fig. 6.7). Specifically, domain I might interact with the CS linker connecting the helices and induce a separation between H1 and H2 to either side of the dimer. The proximity of the extended groove on the surface of domain I (Fig. 6.4) suggests that the small residues of the linker might run through this groove and hold the stem helices H1 and H2 in place. This would involve the disruption of the salt bridge between the linker and H2 (Fig. 6.6). His282 of domain I is located close to this salt bridge, therefore double-protonation of His282 might facilitate the disruption of the salt bridge.

Stiasny et al. [2007] claim to have generated an intermediate of the fusion pathway of a flaviviral E protein by applying alkaline pH, which led to the reversible dissociation of the icosahedral protein envelope. However, the protein was unable to mediate fusion [Stiasny et al. 2007], whereas activation by low pH rapidly leads to fusion. This implies that at high pH the conformational change follows a different pathway than at low pH. More importantly, the reversibility of the alkaline state indicates that this state lay outside the activation pathway, as this is an irreversible process. In contrast, the structural changes observed in the MD simulations in this thesis were triggered by the protonation of histidine residues, i.e. by the modelling of a specifically low-pH dependent effect that would not occur at alkaline pH. Therefore the intermediate configuration from the simulation is likely to represent an activated intermediate of the sE protein.
Model for flaviviral membrane fusion mediated by dimeric E protein

The MD simulations of the sE ectodomain were performed in the absence of the membrane domain, therefore one might expect the structural changes to differ in the full-length E protein in situ the membrane. However, little force is required to bend a membrane [Zimmerberg & Chernomordik 1999], therefore the membrane might follow the structural changes of the protein. The structural changes observed in the simulation of the sE ectodomain after protonation predict increases in the local curvatures of the viral and the host membrane due to interactions between the ectodomain and the membranes at low pH.

The fully conserved residues on the protein surface facing the viral membrane in situ (Fig. 6.6) suggest specific interactions with the stem-anchor region and with the viral membrane. The highly conserved tryptophans and phenylalanines that are found on both surfaces of domain II might interact with the viral and with the target membrane. The hinge motion between domains I and II in the simulation resulted in a curved conformation of the sE dimer (Fig. 6.7A). In situ, interactions between domain I in this conformation, the stem-anchor region and the viral membrane might promote the bending of the viral membrane.

One requirement for fusion in the so-called stalk-pore hypothesis are out-of-plane thermal fluctuations of the bilayers [Zimmerberg & Chernomordik 1999]. In addition, low pH has been reported to result in leakage from liposomes [Drummond et al. 2000], which suggests that low pH increases the thermal fluctuations of the membrane and may thereby promote low-pH-dependent fusion. It is thought that the lipid composition bends the membrane and thereby controls the local concentration and the activity of specific proteins in the membrane [Chernomordik & Kozlov 2003]. Reciprocally, a higher protein concentration, as given on the viral envelope surface, may in turn determine the local lipid composition and membrane bending. In flaviviruses, the stem-anchor regions of both the E and the M protein might facilitate the bending of the viral membrane. The insertion of the fusion peptide might have a similar effect on the target membrane. The membrane domain of M is positioned underneath the base region of domain II (Fig. 5.3). Strong deformations were observed in domain II during the simulation, that may affect the membrane underneath.

In the intermediate configuration from the simulation the fusion peptide was fully exposed (Fig. 5.12B). Exposure of the fusion peptide is required for it to attach to the target membrane, thus the structure fulfills this essential prerequisite for fusion. During the simulation the domains II lifted slightly and exposed the fusion peptide, while the protein remained dimeric and close to its initial flat conformation. In this conformation the dimeric protein might already bind to a target membrane. The buckling motions of the protein domains during the simulation resulted in a conformation in which the “underside”, i.e. the

3 The M protein is cleaved off the N-terminus of the E protein during the maturation of the viral particle and remains in the viral membrane [Li et al. 2008; Zhang et al. 2003]
6.4 Discussion

membrane-facing side, of the fusion region of domain II (orange and yellow respectively) became almost level with the “upside” of domain I (rose) of the respective other subunit (Fig. 6.7A). With a slight increase in the amplitude of the buckling motion beyond what was observed during the 70 ns of simulation, attachment of the membranes to the protein could bring the viral membrane, attached to domain II (yellow or orange), into close proximity with the target membrane, attached to domain I (rose) of the other subunit. When the subunits then separate, the two membranes could come into contact between the subunits (Fig. 6.7B, 5.9) and form a stalk.

Figure 6.8 shows a schematic of the proposed mechanism. Panels A1–3 show cross sections of the viral and the target membrane interacting with the E protein; the sections run parallel to the long principal axis of the E dimer. The black dotted lines indicate a further set of cross sections at 90°, shown in panels B1–3. The membrane domain of E is implied and anchors the ectodomain to the viral membrane. In the figure the E protein ectodomain is sandwiched between the viral and the target membrane (A1–2, B1–2). The construction of the fusion intermediate (A2) was based on the configuration of the ectodomain sE after 70 ns of MD simulation in water. The domain motions observed during the simulation (A2) were extended to obtain the model of a hypothetical active intermediate of the E protein ectodomain (A3, B3). The display of the membranes focuses at two different depths relative to the viewer, indicated by transparent green and pink slabs (A1–3) or boxes (B1–3) respectively. The small inset at the top shows the position of these slabs in bird’s eye view of the viral surface.

The proposed mechanism is as follows: 1) In the pre-fusion conformation at neutral pH, the E protein ectodomain lies flat on the surface of the viral membrane, depicted in Figure 6.8 as a bar beneath the protein; the upper bar represents the target membrane (A1, B1). 2) The structural changes facilitate the insertion of the fusion loops into the target membrane and the bending of the membranes. In the two subunits the structural changes of the dimer and the associated local deformations of the membranes (green and pink slab) move in opposite directions. 3) The E protein in the putative active intermediate conformation (A3, B3). The membranes are brought together between the subunits of the dimer, where they merge and form a stalk, labelled in the figure with a *-symbol. Alternatively the membranes merge between two adjacent dimers, as illustrated by a second stalk in the left part of B3.

In the cryo-EM structure of the mature pre-fusion E protein, the transmembrane helices T1 and T2 of E are located directly underneath the holes between the subunits of the dimer. Thus the protein structure presents no barrier against the possible bending of the viral membrane towards the target membrane by T1 and T2. The fully conserved Phe448, located in the middle of the linker between stem helix H2 and anchor helix T1, might insert into the target membrane and stabilise the putative activated conformation of the protein-membrane...
Figure 6.8 Schematic of the model proposed for flaviviral membrane fusion mediated by dimeric E protein. The dimeric E protein ectodomain is sandwiched in between the viral and the target membrane. A/B1) The protein and the membranes are in a flat conformation. A/B2) Structural changes of the protein promote the bending of the membranes. A/B3) Peak deformations in adjacent sections of the membranes make contact and form fusion stalks, labelled with *-symbols.

A) Cross sections through the viral and the target membrane, parallel to the long principal axis of the E dimer. The membranes are displayed in two slabs, coloured green and pink respectively, and defined by the membrane normal and the orientations of the subunits of the dimer as depicted in the small inset at the top. The green slab lies in front of the pink slab. The model in A2 was constructed using the configuration of the ectodomain sE after 70 ns of MD simulation with all histidines doubly-protonated. B) Cross sections of the models as indicated by black dotted lines in A, perpendicular to the long principal axis of the protein. The transparent green and pink boxes correspond to the membrane slabs shown in A. Filled ovals and circles represent the domains of the E dimer ectodomain: red – domain dII, yellow/orange – dIII, blue – dIII.
complex shown in Figure 6.8AB3. Alternatively, this insertion might perturb the membranes and promote the merger of the membranes.

Based on structural similarities between the fusion proteins of flaviviruses and alphaviruses it was suggested that the proteins are related [Kielian & Rey 2006; Strauss & Strauss 2001]. As there is no sequence identity between flaviviral and alphaviral envelope proteins, the relationship, if any, is distant and the split from a putative common ancestor must have occurred a long time ago. One common functional property among all the low-pH dependent viral envelope proteins known is trimerisation. For instance influenza hemagglutinin and the active forms of alphaviral envelope proteins and class III viral fusion proteins are all trimeric [Backovic & Jardetzky 2009; Gibbons et al. 2004; Weis et al. 1990]. Based on this common property it was proposed that trimerisation is essential for fusion [Earp et al. 2004]. This said, the sequence of events during fusion is not known, and it has been suggested that flaviviral fusion might occur in the absence of trimer formation [Stiasny et al. 2007]. From the simultaneous absence of fusion and trimerisation in mutants [Fritz et al. 2008] it cannot be concluded that fusion requires trimerisation, only that they are correlated. Moreover, a trimeric form of the E protein that mediates fusion implies that there are two different trimeric states, i.e. one active and one inactive. However, only an inactive trimer conformation is known.

In the life cycle of enveloped viruses there are some common environmental constraints which present similar evolutionary constraints on the structures and may have led independently to the parallel evolution of similar structures and fusion mechanisms. Such constraints are for instance the membrane environment and topology before and after fusion, the cell compartment during the biogenesis of the envelope, or the assembly of the viral particle. In the model proposed here for flaviviral fusion, the E protein binds to the target membrane and mediates fusion in the dimeric form, and trimerisation is a consequence of the separation of the dimers, but is not relevant for fusion. This hypothetical mechanism of fusion might have evolved only recently, from an ancestor in which trimerisation was necessary for fusion. Then the flaviviral fusion protein might still display trimerisation after fusion, as a structural atavism from the ancestral protein. Clearly, the model for flaviviral membrane fusion proposed here remains to be validated experimentally. This could be done by cross-linking experiments, e.g. with extended linkers that allow the buckling of the protein and the partial separation of the dimer, but prevent trimerisation.
6.5 Conclusion

A new model is proposed for the low-pH-dependent conformational change of the flaviviral envelope protein ectodomain, in which domain I is released from its meta-stable position between domains II and III. This model is based on a specific fit of crystallographic data and has the advantage that less space is required for the conformational change than previously assumed.

MD simulations of the dengue viral envelope protein ectodomain sE in the pre-fusion, dimeric conformation led to protonation-dependent structural changes that predict interactions between domain I of the ectodomain and the membrane domain of E. This configuration is proposed as an activated intermediate of the ectodomain at low pH and was used for modelling the interactions with the membrane domain and the membrane environment. A model for the mechanism of flaviviral membrane fusion is proposed, in which the deformations of the ectodomain promote bending of the viral and the target membrane and lead to the merger of the membranes and stalk formation. This model differs from previous models in that the active conformation of the E protein is dimeric and sandwiched in between the viral and the target membrane in a relatively flat orientation. Importantly, in the mechanism proposed the viral and the host membrane remain close to each other.

Both the new model of the conformational change and the new mechanism of fusion fit the kinetics of flaviviral membrane fusion, which differ significantly from the kinetics of fusion in other viruses including the alphaviruses. One important conclusion is a different sequence of conformational changes than previously assumed, where the trimerisation step comes after fusion and leads to the inactivation of the protein.