The pH dependence of flavivirus envelope protein structure: insights from molecular dynamics simulations

Carlos Alessandro Fuzo* and Léo Degrève

Grupo de Simulação Molecular, Departamento de Química, Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto-SP, Brazil

Communicated by Ramaswamy H. Sarma

(Received 21 March 2013; accepted 17 July 2013)

The flavivirus membrane fusion is triggered by the acid pH of the endosomes after virus endocytosis. The proposed mechanism involves changes in the protonation state of conserved histidine residues of the E protein present in the viral surface that undergoes a series of structural rearrangements that result in the fusion between the endosome and viral bilayers. We studied the pH dependence of E protein rearrangements of dengue virus type 2, used as a model, in the pH range experimented by the virus along the fusion process. We employed a low computational cost scheme to explore the behavior of the E protein by molecular dynamics (MD) simulations of complete systems that include the protein, the solvent, and ions. The procedure alternates cyclically the update of the ionization states of the protein residues with common MD steps applied to the new ionization configuration. Important pH-dependent protein structure rearrangements consistent with the changes of the protonation states of conserved histidine residues were observed. The involvement of other conserved residues in the flavivirus in the rearrangements was also identified. The results show interesting correlations with a proposed model for the fusion mechanism, as well as the experimentally identified key residues, contributing to a better understanding of the structural changes in protein E that lead to the fusion process.

Keywords: flavivirus; dengue virus; molecular dynamics of envelope protein; pH-dependent protein rearrangements; histidine protonation

Introduction

Dengue virus (DENV) infection has grown dramatically around the world with 40% of the world’s population living in risk areas (World Health Organization, 2012). DENV is a member of the flavivirus genus of Flaviviridae family that includes other significant human pathogens such as west nile virus (WNV), yellow fever virus, tick-borne encephalitis (TBE), and Japanese encephalitis. DENV consists of a nucleocapsid core formed by a single-stranded RNA complexed with multiple copies of the C protein (Becker, 1983). This complex is surrounded by a host-derived lipid bilayer that has two proteins anchored in this outer leaflet: one is a proteolytic residuum of a precursor prM, known as M protein, and the other is the envelope (E) protein with an important role in receptor-binding and fusogenic activities (Harrison, 2008a). The soluble fraction of the E protein (sE) in DENV type 2 (DENV2), which lacks the so-called stem anchor, in the mature virion has been crystallized as a dimer (Modis, Ogata, Clements, & Harrison, 2003). The dimer covers the virion surface with 90 copies in icosahedral scaffold symmetry (Kuhn et al., 2002). Each monomer is a finger-like molecule (Figure 1) consisting of three domains (DI, DII, and DIII).

Flaviviruses infect the host cell through receptor-mediated endocytosis followed by fusion between the endosome and the virus membranes to release the viral RNA into the cytoplasm (Bressanelli et al., 2004; Harrison, 2008a; Modis, Ogata, Clements, & Harrison, 2004). The central event in the viral infection is the fusion mediated by a series of large conformational changes within the E protein (Modis et al., 2003, 2004; Zhang et al., 2004) changing from the prefusion dimer to a postfusion trimer arrangement. The E protein rearrangement is triggered by the low-pH environment inside the endosome resulting in the ensuing steps (Harrison, 2008a, 2008b). The E protein dimers dissociate allowing the monomers to expose the fusion loops to be inserted into the target membrane (Stiasny, Kossel, Lepault, Rey, & Heinz, 2007). At this point, it was postulated that an

*Corresponding author. Email: cafuzo@usp.br

© 2013 Taylor & Francis
extended intermediate of monomers trimerizes and starts
to collapse in a process where DIII folds against DI and
DII, leading to virion and host membrane fusion
(Harrison, 2008a, 2008b). The histidine residues (His),
distributed along all the E protein extension, have been
hypothesized to play a key role in the E protein rear-
rangements because their protonation states change from
uncharged single protonated to positively charged doubly
protonated at the slightly acid pH in the endosomes
(Fritz, Stiasny, & Heinz, 2008; Harrison, 2008b;
Kampmann, Mueller, Mark, Young, & Kobe, 2006;
Mueller et al., 2008; Stiasny, Fritz, Pangerl, & Heinz,
2006). It was concluded, based on the fusion activity of
a series of His to Ala mutants in TBE, that His323
(His317 in DENV2, Figure 1) could be a critical pH sen-
sor together with a possible contribution effect of His146
(His144 in DENV2) where the doubly protonated forms
may destabilize the DI–DIII interface (Fritz et al., 2008).
In the same study, a double His to Ala mutation of
His248 and His287 (His244 and His282 in DENV2)
impaired the fusion to the same extension as the one
obtained by a His323 mutation. The effect of this double
mutation on the fusion mechanism is not completely
understood, but it was hypothesized that the double
mutation impairs one of the steps of the fusion reaction
such as the transition from the intermediate extended E
protein structure to the active trimer (Harrison, 2008b).
Furthermore, the replacement of individual His with a
nonionizable amino acid in WNV revealed that a single
His is not required as a pH sensor suggesting that other
features of the E protein must be involved in the fusion
process (Nelson, Poddar, Lin, & Pierson, 2009).
Computational chemistry has been used to better
understand the molecular mechanisms involving biologi-
cal processes. In this sense, methods like molecular
dynamics (MD) have great potential to contribute to
understand the behavior of E protein. The MD method is
a powerful and widely used technique for understanding
the dynamic and structural features of proteins (Durrant
& McCammon, 2011; Karplus & Kuriyan, 2005). In the
case of DENV2, MD was used to study the influence of
the His residues protonation on the E protein structure
(Degrève, Fuzo, & Caliri, 2012; Dubey, Chaubey, &
Ojha, 2011; Kampmann et al., 2006; Mueller et al., 2008;
Prakash, Barducci & Parrinello, 2010). The understanding
of the conformational E protein changes at an atomic
level helps in designing and targeting antiviral agents
(Kaufmann & Rossmann, 2011). In the present work, the
behavior of the monomer of the DENV2 E protein as a
function of the pH was studied by MD particularly in the
pH range that has been attributed to trigger the structural
changes that lead to the fusion process.

Models and methods
The pH of the medium commonly applied in MD studies
of proteins is generally modeled by assigning the
protonation state of the ionizable groups based on the pK_a
values obtained generally by direct titration in aqueous
solutions or by the pK_a calculated in the protein initial
structure. This can be done with the aid of many methods
(Alexov et al., 2011). This scheme is an accepted approp-
riate approach because it takes into account the effects
of the local protein structure of the protonable groups on
their pK_a. However, the effects caused by the protein
structure changes on the pK_a are not considered along the
MD calculations that are performed in systems with fro-
zen ionization. Constant-pH MD simulation algorithms,
where the protonation states can be updated, have been
proposed to overcome this problem (for a review see
Alexov et al., 2011). Some of the most popular algorithms
employ implicit solvent models (Lee, Salsbury, & Brooks,
2004; Mongan, Case, & McCammon, 2004). Other meth-
ods have also been developed to treat the solvent explic-
itly (Baptista, Teixeira, & Soares, 2002; Donnini, Tegeler,
Groenhof, & Grubmüller, 2011; Wallace & Shen, 2011).
The solvent plays an important role in the maintenance of
protein structures. For this reason, it is interesting to con-
sider explicitly the solvent in MD studies on the behavior
of protein structures at defined pH. However, the great
computational cost makes explicitly solvent algorithms

frequently inaccessible for the study of large proteins such as the E protein. We employed a lower computational cost scheme to explore the behavior of the E protein as a function of pH by explicit solvent MD. The scheme, illustrated in Figure 2, consists of cycles of short MD simulation sets (MDS) carried out at a defined pH, where an initial assignment of the ionization state of each ionizable site $i$ that is updated based on the $pK_a$ values obtained with the application on the initial MDS configuration of the PROPKA method (Olsson, Sondergard, Rostkowski, & Jensen, 2011). The protonation states of all sites are updated using the new $pK_a$ values and assigning the ionizable state if the condition $\text{pH} \leq pK_a$ is satisfied. If the former condition is not satisfied to His, one of the $\delta$ or $\epsilon$ tautomeric forms (the protonation of the ND1 or the NE2 atom) is attributed with the same probabilities. The next step is a short MD of $0.2$ ps using a $1$ fs time step conducted to consider the solvent relaxation but harmonically restraining the positions of all atoms, with the exception of the hydrogen atoms, of the protein using a $1000 \text{kJ mol}^{-1} \text{nm}^{-2}$ force constant. The last MDS phase is a $20$ ps MD performed with a $2$ fs time step for the adaptation of the protein to the new protonation states producing a new protein structure that will be the input for the next MDS. This time interval of $20$ ps is also long enough to equilibrate the water molecules. The entire MDS procedure is repeated $N$ times. This scheme was applied to the study of the E protein monomer at various pHs beginning with slightly basic (pH 8) up to sufficiently acidic (pH 5) media allowing the side chain of His residues to change from unprotonated to protonated.

The initial configuration of the protein was obtained from chain A of the dimer under the pdb code 1OKE (Modis et al., 2003). The protonation states, at pH 7, were attributed based on the bulk solvent $pK_a$. The initial system was constructed by inserting the E protein monomer of the DENV2 in the center of an $8.0 \times 9.0 \times 16.0 \text{nm}^3$ box which was supplemented with water molecules (about 36,000) modeled by the SPC/E model (Berendsen, Grigera, & Straatsma, 1987), together with the appropriate number of ions to neutralize the system and to reach an ionic strength equal to $150 \text{mM}$. The system was then submitted to energy minimization using a steepest descent algorithm to eliminate undesirable forces. Then, six simulations of $10$ ps each were performed, increasing the temperature along the sequence $50, 100, 150, 200, 250$, and $300 \text{K}$. Finally, simulations at pH 5, 6, 6.5, 7, and 8 were performed for $50$ ns, each applying the proposed $pK_a$ determination scheme (Figure 2). The simulations were carried out using the GROMACS 4.5.1 simulation package (van der Spoel et al. 2005) with the GROMOS96 43A1 force field (van Gunsteren et al., 1996). The protein covalent bonds involving hydrogen atoms were constrained by the LINCS algorithm (Hess, Bekker, Berendsen, & Fraaije, 1997), while the SETTLE algorithm (Miyamoto & Kollman, 1992) was used to maintain the water molecules according to the rigid SPC/E model. The temperature of the solvent was regulated by the Berendsen algorithm (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984) at $300 \text{K}$, and a minimal invasive thermostat was applied to the protein (Lingenheil, Denschlag, Reichold, & Tavan, 2008). A twin-range cut-off with short- and long-range cut-offs equal to $0.9$ and $1.4 \text{nm}$, respectively, was applied to the interactions, and the neighbor’s pair list was updated every $10$ time steps. The electrostatic interactions were calculated with the reaction field with infinite permittivity outside the cut-off (van der Spoel, Lindahl, Hess, van Buuren et al., 2005). The MD motion equations were integrated by the leap-frog algorithm (Hockney, Goel, & Eastwood, 1974). At each pH, $2500$ MDS of $20$ ps were performed, resulting in a total of $50$ ns of protein conformational sample. The coordinates of all the atoms were recorded at every cycle (every $20$ ps) for further analysis conducted with the GROMACS tools (van der Spoel et al., 2005), Pymol (The PyMOL Molecular Graphics System), and in-house programs. The fraction of protonated state (s) for each protonable group in function of pH was used to calculate the $pK_a$ after the fitting to the Hill equation:

\[
s = \frac{1}{1 + 10^{[\text{pH}-pK_a]}}
\]

where $n$ is the Hill coefficient that is also obtained with the fit.

**Results and discussion
Evaluation of the protonation states**

The protonation states of the His, Lys, Arg, Glu, and Asp residues and of the N and C-terminus were investigated all along the simulations. However, the His
residues were the main originator for the protein protonation changes with all the His residues gradually going from a single, at pH 8, to a double, at pH 5, protonation state. The titration of the His residues was computed along all the MDS as the fraction of the doubly protonated forms (Figure 3). The resulting profile is characteristic of a titration curve, as was already observed in recent protein MD simulations conducted at constant pH (Machuqueiro & Baptista, 2006, 2007; Yang, Zhang, Ho, & Ding, 2011). The fraction of doubly protonated residues decreased from about 1.0 at pH 5 to 0 at pH 8. The pH at half protonation is $6.61 \pm 0.04$ in agreement with the reference $pK_a$ of 6.50 used by PROP-KA for the His side chain (Olsson et al., 2011). The $pK_a$ of the individual His residues, listed in Table 1, was calculated from the half-protonation state for each His residue. The individual His $pK_a$ values range from 5.96 ± 0.15 to 7.50 ± 0.11 with some residues presenting $pK_a$ values near the reference value ($pK_a = 6.50$): this is the case for His27, His144, His244, His261, His282, and His346. Some $pK_a$ are larger than the reference, the case for His94, His149, His209, and His317, or smaller as observed with His158. These disturbances are correlated with the location of the residue in the protein structure and with its vicinity explaining why the structural rearrangements consequent of the pH changes result in different $pK_a$. Another group that presents a modification in its protonation state along the 5–8 pH range is the N-terminus (Figure S1) since its calculated $pK_a$ is 6.71 ± 0.02. Unique other partial protonations were observed for Asp42, Glu44, Glu147, Glu338, and Glu370 at pH 5 and 6 (Figure S2 and Table S1). These results indicate that the protonation states along the pH range allow to investigate the pH behavior of the E protein. The long simulation time of 50 ns in each case is sufficient for observing the first steps of the structural protein arrangements occurring in the function of the pH.

### Behavior of E protein structure as a function of pH

The main interest in knowledge of the behavior of the E protein structure in the function of pH is the protein’s essential role in the fusion process with structural changes triggered by variations in His residue protonation (Fritz et al., 2008; Harrison, 2008b; Kampmann et al., 2006; Mueller et al., 2008; Stiasny et al., 2011). The protein experiences structural changes when the experimental E protein structure is put in an aqueous medium where the ionic strength is 150 mM. This behavior can be a consequence of a series of factors such as the loss of contacts with the other monomer, once the monomer used was extracted from the dimer structure, the relaxation of the crystallographic structure in the aqueous medium by direct influence of the solvent and of the ions, and the protonation effects. The major structure changes occur during the first 20 ns as can be observed by the $C_\alpha$ root-mean-square deviations (RMSDs) in relation to the average structures (Figure 4).

![Figure 3. The fraction of protonated His residues along the 5–8 pH range. The solid line is the fit of the Hill equation. The $pK_a$ at the inflection point is $6.61 \pm 0.04$.](image3)

Table 1. The individual $pK_a$ of the His residues calculated after the Hill equation fitting to the fraction of the protonated states along the simulation.

| Residue | $pK_a$ |
|---------|--------|
| His27   | 6.35 ± 0.03 |
| His94   | 7.50 ± 0.11 |
| His144  | 6.29 ± 0.08 |
| His149  | 6.78 ± 0.01 |
| His158  | 5.96 ± 0.15 |
| His209  | 7.21 ± 0.08 |
| His244  | 6.68 ± 0.11 |
| His261  | 6.60 ± 0.03 |
| His282  | 6.54 ± 0.19 |
| His317  | 6.98 ± 0.01 |
| His346  | 6.66 ± 0.05 |

![Figure 4. Overall protein structure deviation from the averaged structures obtained along the entire simulation time for all pH values.](image4)
The RMSD values are greater than .35 nm at the beginning of the simulations, decreasing to values near .22 nm at about 20 ns, and remaining at this level up to 50 ns, thus indicating that the simulation times were sufficient for the convergence to representative structures at each pH. The exception is the simulation at pH 7 in which the structures undergo some changes at 25 and 27 ns when the RMSDs reach values near .42 nm before oscillating near the values of the averaged structures presenting fluctuations larger than those observed at the other pHs. This is a consequence of the structure disturbances due to the small fraction (.3) and the lack of persistence of the His protonation (Figures 3 and S1). Consequently, the first 30 ns of the simulation were considered the equilibration phase, and the protein averaged structures for each pH, which will be used in the next analysis, were, therefore, calculated in the 30 to 50 ns interval. It is interesting to observe that the calculated RMSDs indicate that the structures converge quickly despite their large fluctuations due to the predefined pH changes, as it will be seen below, thereby indicating that the energy barriers between local minima can be easily overcome when the protonation states are continuously updated as it must occur in the real systems.

A bend about a hinge between DI and DII is essential to the viral life cycle (Bressanelli et al., 2004; Modis et al., 2004; Zhang et al., 2004). To observe if such a bend can be pH dependent, the hinge angles ($\varphi$) were calculated relative to the crystallographic structure (Figure 5) fitting DI on the experimental structure. The first conclusion is that the great rearrangement of DII relative to DI is the bend about the hinge region. The E protein adopts an extended structure when the pH becomes slightly acid since $\varphi$ reaches $-43.5^\circ$ at pH 6.5. This conclusion is interesting because the existence of an extended intermediate structure has been postulated in the proposed fusion mechanism mediated by the E protein (Harrison, 2008a, 2008b). Determining the responsible residues that lead to this extended structure of the E protein will lead to major knowledge of the fusion process and of its mechanisms. Another interesting point is that the extended structure is observed only at the slightly acid pH 6.5 where the protonation fraction of the His residues is about .5. Thereafter, in more acid media at pHs 6 and 5, the E protein adopts more compact structures with $\varphi$ equal to 7.3$^\circ$ and 17.9$^\circ$. This behavior is also observed in early studies of the E protein simulated with all the His residues in double protonated conditions where a more compact structure was observed due to the large flexibility of the hinge region at physiological conditions (Kaufmann & Rossmann, 2011).

The destabilization of the DI–DIII interface has also been considered important in the fusion mechanism (Bressanelli et al., 2004; Fritz, Stiasny, & Heinz, 2008; Modis et al., 2004; Stiasny et al., 2011). No clear feature like a DI–DII hinge motion was observed for DIII fitting DI on the experimental structure. To observe the direction of the motion of DIII, the displacement of the residues inside the DIII was computed in the averaged structures using the experimental structure as the reference after fitting DI (Figure 6). The movement of DIII relatively to DI shows that the main drift is a global rotation of angle $\lambda$ around the center formed by the residues presenting the smallest displacement (Figure 6). An exception is observed at pH 6 where the shifts occur only in one direction. In the other cases, the rotations are

Figure 5. Main reorganization in the protein E structure in the function of the pH. The changes in the hinge angle are in relation to the experimental structure of a monomer extracted from the dimer structure 1OKE. The DI of the 30–50 ns averaged structures was superimposed to the DI of chain A of 1OKE. For better visualization, only the DII at each pH and all the domains of 1OKE are shown. The hinge region of 1OKE is also identified. The angle $\varphi$ is defined as the angle between the segments connecting the center of mass of the hinge to the center of mass of DII in 1OKE and in the averaged structures produced at different pH.
The rotation occurs in the opposite direction at pH 5 with $\lambda$ equal to $-54.3^\circ$. The pH 6.5 structure is extended due to the changes at the DII–DIII hinge ($\varphi = -43.5^\circ$). Changes at the DI–DIII interface were also observed resulting in DIII spinning in one direction at pH 6.5 and in the other direction in more acid conditions. These results demonstrate that the primary response of the E protein monomer to the acidification of the surrounding medium is the formation of an extended structure with changes at the DI–DII and DI–DIII interfaces. However, at pHs 5 and 6, the settlement of the DII is going in the opposite direction so that the E protein adopts a more compact structure. The more extended structure of the E protein is found at pH 6.5 and the most compact structure at pH 6. The pH range between these extremes is very sharp, emerging close to the His $pK_a$ of 6.5, which indicates that the E protein flexibility is very sensitive to the His protonation.

Local moves are responsible for the pH-dependent E protein rearrangements

The RMSD per residue ($\text{RMSD}_{\text{res}}$) were calculated for each domain by fitting their conserved regions in relation to the domains averaged structures at pH 8 (Figure 7) to identify the key residues, or the key regions, that may be altered by the pH. The structure at pH 8 was used as reference. An alternative is the crystalline structure that is quite different from the relaxed structures produced by MD. The profiles of the $\text{RMSD}_{\text{res}}$ present similar behaviors at all pHs with the exception of pH 7 where the $\text{RMSD}_{\text{res}}$ values are due to the small fraction of the His protonation and to the lack of its time persistence. The residues of DII with the largest $\text{RMSD}_{\text{res}}$ are located in the tip of the E protein, the region of the fusion peptide responsible for the insertion into the host membrane according to the proposed fusion mechanism (Bressanelli et al., 2004; Harrison, 2008a; Modis et al., 2004). The RMSDs of this region are higher than the observed in the stable regions since they are equal to .27 nm at pHs 6 and 6.5, and .34 nm at pHs 5 and 7, with the $\text{RMSD}_{\text{res}}$ reaching 1.59 nm in the pH 5 case. It was observed early that the rearrangements of the DII and DIII occur with respect to DI; thus, the major changes in DI, denoted by the largest $\text{RMSD}_{\text{res}}$ involve mainly the residues near the DI–DII and DI–DIII interfaces. Therefore, a detailed analysis of these interfaces in the function of the pH will target the key residues in the structural rearrangements of the E protein.

Changes at the DI–DII interface

Four segments present at the DI–DII interface connect DI and DII (Butrapet et al., 2011). They are designated by H1 (Lys51–Ala54), H2 (Val130–Leu135), H3 (Ser186–Met196), and H4 (Thr265–Leu283). Comparing H1, H2, H3, and H4 shows that the largest $\text{RMSD}_{\text{res}}$ are observed in H2 and in H3. Due to the proximity between Gln167 and H3, the repositioning of Gln167 produces $\text{RMSD}_{\text{res}}$ peaks in the pHs 6.5 and 7 cases reaching 1.22 nm at pH 6.5. This indicates that the large rearrangements of the E protein at pH 6.5, which were detected mainly through $\varphi$, are due to the rearrangements in H3 and in its neighborhoods. The H3 segment undergoes a torsion that shortens its main chain from about 1.90 nm, in the experimental structure, to 1.15 nm at pH 6.5 (Figure 8 and Table 2). Slight decreases in the H3 sizes to 1.67 and 1.65 nm are also observed at pHs 7 and 8, while in the acid pHs 5
and 6, the opposite behavior is observed with the H3 largest sizes (1.99 and 2.03 nm). The fluctuations of \(\phi\) and \(d\) (Figure 8c) at pH 6.5 indicate that the packing of H3 is to be correlated with the approach between His209 and His282 that causes strains sufficient to translate the entire DII domain.

The electrostatic interactions play a key role in the overall behavior of the E protein depending on the pH. At pH 6.5, the population of the double protonated form of His209 (Figure 8c) is high (\(pK_a\) of 7.21 ± .08) favoring the approach of His209 to H3, where the Asp192 residue is negatively charged. Moreover, the \(pK_a\) of His282 (6.54 ± .19), lower than the \(pK_a\) of His209, is found in the pH range where the protein experiences the largest changes in its structure, which is transformed from an extended to the most compact one between pH 6.5 and pH 6. This remark indicates that when the His282 residue is also doubly protonated, the His209 and His282 residues tend to remain apart just as observed at more acid pHs due to the electrostatic repulsion of their positive charges. The fraction of protonated His209 and His282 at pH 7 is compatible with the fraction observed at pH 6.5, meanwhile the changes in the protonation states along the time are different. At pH 7, the doubly protonated states of His209 and His282 occur simultaneously during some periods (Figure S3) that are correlated with the largest RMSD fluctuations (Figure 4). At pH 7, many transitions between the protonated and nonprotonated states (Figure S1) that induce fluctuations of the protein structure located mainly in DII around His244. At pH 6.5, the double protonation of His244, that is located in a turn, induces the formation of a salt bridge with Asp98 (Figure S4), the last residue of a \(\beta\) sheet. This salt bridge stabilizes the nearby local structures of DII promoting the structure alterations observed in H3 and in its neighborhood.

**Changes at the DI–DIII interface**

The rearrangements in DIII stem from the protonation variations of the key residues N-terminus, His144, Glu147, and His317. The changes in their neighborhood are detected by the shortest distances and by the structures of the contact region between DI and DIII as shown, respectively, in Table 3 and Figure 9. The changes in the hinge region between DI and DIII at pH 6.5 are caused by the protonation of the N-terminus that favors the approximation of the negatively charged Glu368 residue to the positive charges of the N-terminus. Simultaneously, the distance between His317 and DI increases promoting the rotation of DIII relatively to DI in the direction observed earlier (Figure 6). The positively charged residues His144 and N-terminus at pH 6 (Figure S1) are maintained close to Arg9 due to the presence of the negatively charged Asp42, Glu147, and Glu368 residues present in their vicinity. The displacements of Glu147 and Glu368 in the direction of Arg9 and His144 and the N-terminus residues cause tensions able to shift entirely the relative position of DIII (Figure 6). The protonation of Glu147 at pH 5 causes the disturbances in the pH 6 configuration of the N-terminus, Asp42, and also His144. Therefore, Glu147 and the N-terminus residues are jointly removed from the vicinity of His144 that remains close to Asp42 and to Glu368. This rearrangement induces the large rotation of DIII in the direction opposed to the observed at neutral and slightly acid pH, placing His317 close to the negatively charged region of DI, i.e. close the negatively charged residues Asp10 and Glu13. It is interesting to observe that the Ser145–Glu148 segment of DI remains stable only relatively to DIII but not to DI (Figure S5), thereby showing that the changes in the DIII orientations are not dissociated from the internal movements of DI. Furthermore, the His144 residue, that has been

![RMSD by fitting the stable regions with reference to the averaged pH 8 structure (DI: segments Asp22–Lys47, Tyr137–Ile141, Lys160–Glu184, and Gly281–Gly296; DII: Pro53–Lys64, Thr120–Pro132, Asn194–Ile232, and Gly254–280; DIII: all the residues). The domains are identified by the red (DI), yellow (DII), and blue (DIII) colors. The position of the His residues are indicated as open circles. The four segments located at the DI–DII interface are identified as H1 (Lys51–Ala54), H2 (Val130–Leu135), H3 (Ser186–Met196), and H4 (Thr265–Leu283).](image-url)
Figure 8. Initial (a) and final (b) structure obtained at pH 6.5 highlighting the His209 and His282 residues as spheres and the H3 residues as sticks. The residue His209 is located in DII at the beginning of an α-helix (segment His209–Asp215). At pH 6.5, His209 is located close to the backbone oxygen atoms of the residues Arg188, Gly190, and Ser192 and close to one oxygen atom of the Ser192 side chain in the twisted H3 loop. The His282 residue is located near the other face of the H3 loop where the backbone oxygen atoms of Pro187, Thr189, Leu191, and the other oxygen of the side chain of Ser192 are present. (c) The hinge angle $\varphi$, the distance $d$ between the center of mass of the His209 and His282 side chains, the (1 or 0) full or not full protonation (1 or 0) of His209 and His282, and the sum of these last data at pH 6.5.
considered important in the fusion mechanism (Fritz et al., 2008; Harrison, 2008b; Kampmann et al., 2006; Mueller et al., 2008; Stiasny et al., 2011), is the hinge region of the DIII rearrangements.

Table 2. Distances (nm) between the $C_\alpha$ atoms of the Ser186 and Met196 and the minimum distance between the heavy atoms of His209 and His282 in the averaged structures.

|           | pH 5 | pH 6 | pH 6.5 | pH 7 | pH 8 |
|-----------|------|------|--------|------|------|
| Ser186–Met196 | 1.99 | 2.03 | 1.15   | 1.67 | 1.65 |
| His209–His282 | 1.61 | 1.71 | .96    | 1.76 | 1.57 |

Table 3. Shortest distances (nm) between the heavy atoms of the side chains of the residues, or of the segments, listed in the first column (where Nter is the N-terminus) in the averaged structures at different pH and in chain A of 1OKE.

| Pair          | pH 5 | pH 6 | pH 6.5 | pH 7 | pH 8 | 1OKE |
|---------------|------|------|--------|------|------|------|
| His317–DI     | .27  | .89  | .54    | .58  | .34  | .31  |
| His317–Arg9   | .66  | .85  | .92    | .64  | .42  | .37  |
| Arg9–Glu368   | .31  | .32  | .77    | .41  | .49  | .28  |
| Arg9–Asp42    | .69  | .44  | .62    | .36  | .44  | .42  |
| Nter–Arg9     | 1.13 | .51  | .66    | .56  | .52  | .71  |
| Nter–Asp42    | .48  | .48  | .35    | .40  | .32  | .31  |
| Nter–His144   | .64  | .38  | .34    | .40  | .30  | .37  |
| Nter–Glu147   | .42  | .46  | 1.36   | 1.02 | 1.30 | .94  |
| Nter–Glu368   | .91  | .31  | .33    | .64  | .67  | .74  |
| His144–Asp42  | .27  | .32  | .33    | .50  | .42  | .31  |
| His144–Glu147 | .77  | .33  | 1.05   | 1.03 | 1.12 | 1.22 |
| His144–Glu368 | .27  | .45  | .50    | .38  | .56  | .34  |

Figure 9. The structure of the contact region between DI (red) and DIII (blue) after the fit of DI in the experimental structure showing the key residues involved in the DIII changes in function of pH. The structures at each pH are the ones with the minor RMSDs in relation to the average structure.

Role of the conserved residues in E protein rearrangements as a function of pH

Conservation of the key residues, responsible for the pH-dependent rearrangements of the DII and DIII observed
early, was analyzed from the Pfam seed alignment (Punta et al., 2012) with the entries PF00869, which correspond to the central and dimerization domains (DI and DII), and PF02832, immunoglobulin-like domain (DIII), of the flavivirus.

The former observations indicate the important role of the residues Asp98, Asp192, His209, His244, and His282 in the E protein rearrangements originating at the hinge region DI–DII since the protein assumes an extended structure only at pH 6.5 due to the delicate combination of protonation changes and rearrangements of local structures. The key residues involved in the DI–DII hinge motion were searched in the seed alignment presenting the following percentage of conservation (in parentheses): Asp98 (97%), Asp192 (82%), His209 (65%), His244 (94%), and His282 (97%). With the exception of His209, these data show that these residues are highly conserved. In a previous study, a double His to Ala mutation in the corresponding His244 and His282 in TBE demonstrated that these conserved residues are important for virus fusion (Fritz et al., 2008). Although effect of this double mutation on the fusion mechanism is not completely understood, they have been hypothesized to affect the transition from the extended intermediate structure of the monomer to the folded trimer during the fusion reaction mechanism (Harrison, 2008b). The extended structure observed at pH 6.5 is consistent with such a hypothesis suggesting that the Asp98 and Asp192 residues also play an important role in forming and stabilizing the extended structure. The His209 residue presents middle conservation of 65%, indicating that this residue may not be essential for E protein function. However, the next neighbor of the His209 residue is a positively charged residue (Arg or Lys) with 85% conservation. Interestingly, one of these positively charged residues is always present in the species when His209 is absent. This indicates that one of them can take the place of His209 in the stabilization of the local structures observed early in the hinge region of DI–DII. The His144 and His317 residues, conserved with 97 and 100% of occurrences, are of great significance in the fusion process (Fritz et al., 2008). Despite the importance of the His144 and His317 residues, the present results have shown that the N-terminus and the Arg9 (97%), Asp42 (97%), Glu147 (11%), and Glu368 (100%) residues are also required for the rearrangements of DIII. These data reveal that the key residues involved in the DI–DIII rearrangements are highly conserved in the flavivirus with the exception of Glu147. However, the Glu147 residue is located in a segment (Ser145–His158) where the sequences and sizes are not conserved. Another fact should draw attention for being curious and quite significant: negatively charged residues (Glu and Asp) that can play the same role as Glu147 are recognized in the seed alignment of this segment in 97% of the species. Despite the presence of the His149 and His158 residues in the Ser145–His158 segment, no correlation between the changes in the protonation of these residues and the structural rearrangements observed in DIII was found. The analysis of the alignment has shown that these residues are not highly conserved and that the occurrence of at least one His residue in the Ser145–His158 segment is 60%. The conclusion of this last analysis is that the key residues, responsible for the rearrangements of DII and DIII in the E protein observed in the simulations, are extremely well conserved in the flavivirus.

Conclusions

The data produced by the simple procedure to actualize the ionization state of the residues show modifications in the overall dengue E protein structure in function of pH. Structural changes in the E protein are caused mainly by the changes in the protonation states of the His residues that are considered important for triggering the structural changes that lead to the fusion process. In conjunction with His, changes in protonation or the presence of other residues, which are also conserved in flavivirus, also contribute to structural changes of E protein observed in the pH range studied. The information obtained from these studies contributes to a better understanding of the fusion process mediated by the E protein. The application of a simple scheme, compared with the computationally very expensive and more robust algorithms to mimic the protonation changes as a function of pH, allows to study large protein like the E protein producing information that are consistent with the experimental data. The great advantage is that good insights of the structural changes of large proteins, depending on the pH, can be achieved. Nevertheless, the role of the His residues, as well as of the other residues in triggering the fusion process, is still a matter of debate requiring extensive studies to elucidate completely all points of the fusion process that is extremely complex.

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2013.827132.

Acknowledgements

The authors are thankful to Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support.

References

Alexov, E., Mehler, E. L., Baker, N., Baptista, A. M., Huang, Y., Milletti, F., ... Word, J. M. (2011). Progress in the prediction of pKa values in proteins. Proteins, 79, 3260–3275.
Baptista, A. M., Teixeira, V. H., & Soares, C. M. (2002). Constant-pH molecular dynamics using stochastic titration. *Journal of Chemical Physics*, 117, 4184–4200.

Becker, Y. (1983). *Molecular virology: Molecular and medical aspects of disease-causing viruses in man and animals*. The Hague: Martinus Nijhoff.

Berendsen, H. J. C., Grigera, J. R., & Straatsma, T. P. (1987). The missing term in effective pair potentials. *The Journal of Physical Chemistry*, 91, 6269–6271.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*, 81, 3684–3690.

Bressanelli, S., Stiasny, K., Allison, S. L., Stura, E. A., Duquerroy, S., Lescar, J., ..., Rey, F. (2004). Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. *EMBO Journal*, 23, 728–738.

Butetrap, S., Childers, T., Moss, K. J., Erb, S. M., Luy, B. E., Calvert, A. E., ..., Huang, C. Y. H. (2011). Amino acid changes within the E protein hinge region that affect dengue virus type 2 infectivity and fusion. *Virology*, 413, 118–127.

Degrève, L., Fuzo, C. A., & Caliri, A. (2012). Extensive structural change of the envelope protein of dengue virus induced by a tuned ionic strength: Conformational and energetic analyses. *Journal of Computer-Aided Molecular Design*, 26, 1311–1325.

Donnini, S., Tegeler, F., Groenhof, G., & Grubmüller, H. (2011). Constant pH molecular dynamics in explicit solvent with λ-dynamics. *Journal of Chemical Theory & Computation*, 7, 1962–1978.

Dubey, K. D., Chaubey, A. K., & Ojha, R. P. (2011). Role of pH on dimeric interactions for DENV envelope protein: An insight from molecular dynamics study. *Biochimica et Biophysica Acta*, 1814, 1796–1801.

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, 9, 71–79.

Fritz, F., Stiasny, K., & Heinz, F. X. (2008). Identification of specific histidines as pH sensors in flavivirus membrane fusion. *The Journal of Cell Biology*, 183, 353–361.

Hockney, R. W., Goel, S. P., & Eastwood, J. (1974). Quiet high-resolution computer models of plasma. *Journal of Computational Physics*, 14, 148–158.

Kampmann, T., Mueller, D. S., Mark, A. E., Young, P. R., & Kobe, B. (2006). The role of histidine residues in low-pH-mediated viral membrane fusion. *Structure*, 14, 1481–1487.

Karpplus, M., & Kurijan, Y. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences*, 102, 6679–6685.

Kaufmann, B., & Rossmann, M. G. (2011). Molecular mechanisms involved in the early steps of flavivirus cell entry. *Microbes & Infection*, 13, 1–9.

Kuhn, R. G., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lencehes, E., ..., Strauss, J. H. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell*, 108, 717–725.

Lee, M. S., Salsbury, F. R., & Brooks, C. L. (2004). Constant-pH molecular dynamics using continuous titration coordinates. *Proteins*, 56, 738–752.

Lingenheil, M., Denschlag, R., Reichold, R., & Tavan, P. (2008). The “hot-solvent/cold-solute” problem revisited. *Journal of Chemical Theory & Computation*, 4, 1293–1306.

Machuqueiro, M., & Baptista, A. M. (2006). Constant-pH molecular dynamics with ionic strength effects: Protonation-conformation coupling in decalysine. *The Journal of Physical Chemistry B*, 110, 2927–2933.

Machuqueiro, M., & Baptista, A. M. (2007). The pH dependent conformational states of kyotorphin: A constant-pH molecular dynamics study. *Biophysical Journal*, 92, 1836–1845.

Miyamoto, S., & Kollman, P. A. (1992). SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. *Journal of Computational Chemistry*, 13, 952–962.

Modis, Y., Ogata, S., Clements, D., & Harrison, S. C. (2003). A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proceedings of the National Academy of Sciences*, 100, 6986–6991.

Modis, Y., Ogata, S., Clements, D., & Harrison, S. C. (2004). Structure of the dengue virus envelope protein after membrane fusion. *Nature*, 427, 313–319.

Mongan, J., Case, D. A., & McCammon, J. A. (2004). Constant pH molecular dynamics in generalized born implicit solvent. *Journal of Computational Chemistry*, 25, 2038–2048.

Mueller, D. S., Kampmann, T., Yennamalli, R., Young, P. R., Kobe, B., & Mark, A. E. (2008). Histidine protonation and the activation of viral fusion proteins. *Biochemical Society Transactions*, 36, 43–45.

Nelson, S., Poddar, S., Lin, T. Y., & Pierson, T. C. (2009). Protonation of individual histidine residues is not required for the pH-dependent entry of west nile virus: Evaluation of the “histidine switch” hypothesis. *Journal of Virology*, 83, 12631–12635.

Olsson, M. H. M., Sondergard, C. R., Rostkowski, M., & Jensen, J. H. (2011). PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory & Computation*, 7, 525–537.

Prakash, M. K., Barducci, A., & Parrinello, M. (2010). Probing the mechanism of pH-induced large-scale conformational changes in dengue virus envelope protein using atomistic simulations. *Biophysical Journal*, 99, 588–594.

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., ..., Finn, R. D. (2012). The Pfam protein family database. *Nucleic Acids Research*, 40, D290–D301.

Stiasny, K., Fritz, R., Pangerl, K., & Heinz, F. X. (2011). Molecular mechanisms of flavivirus membrane fusion. *Amino Acids*, 41, 1159–1163.

Stiasny, K., Kossel, C., Lepault, J., Rey, F. A., & Heinz, F. X. (2007). Characterization of a structural intermediate of flavivirus membrane fusion. *PLoS Pathogens*, 3, 191–199.

The PyMOL Molecular Graphics System (Version 1.5.0.4). Schrödinger, LLC.

van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26, 1701–1718.

van der Spoel, D., Lindahl, E., Hess, B., van Buuren, A. R., Apol, E., Meulenhoff, P. J., ..., Berendsen, H. J. C. (2005). Gromacs user manual (Version 4.5). Retrieved from http://www.gromacs.org

van Gunsteren, W. F., Biller, S. R., Eising, A. A., Hunenberger, P. H., Kruger, P., Mark, A. E., ..., Tironi, I. G. (1996). *Biomolecular simulation: the GROMOS96 manual and user guide*. Zürich: vdf Hochschulverlag AG an der ETH Zürich.
Wallace, J. A., & Shen, J. K. (2011). Continuous constant pH molecular dynamics in explicit solvent with pH-based replica exchange. *Journal of Chemical Theory & Computation, 7*, 2617–2629.

World Health Organization. (2012). Dengue and severe dengue. Fact sheet 117. Retrieved from http://www.who.int/mediacentre/factsheets/fs117/en/

Yang, L., Zhang, J., Ho, B., & Ding, J. L. (2011). Histidine-mediated pH-sensitive regulation of M-Ficolin: GlcNAc binding activity in innate immunity examined by molecular dynamics simulations. *PLoS ONE, 6*, 1–9.

Zhang, Y., Zhang, W., Ogata, S., Clements, D., Strauss, J. H., Baker, T. S., … Rossmann, M. G. (2004). Conformational changes of the flavivirus E glycoprotein. *Structure, 12*, 1607–1618.