

# Structure and dynamics of the monomer of protein E of dengue virus type 2 with unprotonated histidine residues

L. Degrève and C.A. Fuzo

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

Corresponding author: L. Degrève  
E-mail: leo@fcfrp.usp.br

Genet. Mol. Res. 12 (1): 348-359 (2013)  
Received February 16, 2012  
Accepted October 26, 2012  
Published February 7, 2013  
DOI <http://dx.doi.org/10.4238/2013.February.7.5>

**ABSTRACT.** The surface of the dengue virus is composed of 180 copies of a multifunctional envelope glycoprotein that acts at several stages of viral infection. When the virus is in the endosome, these glycoproteins undergo major conformational rearrangements owing to the protonation of histidine side chains. This protonation allows for the formation of trimers, thereby triggering fusion between the viral and the host membranes. In this study, we examined the behavior of a monomer of this key protein containing unprotonated histidine side chains before the stage of trimer formation using explicit solvent molecular dynamics at various ionic strengths. The extended secondary structures, which contribute to protein stabilization, are smaller than those observed in a previous study involving monomers containing the protonated histidine. However, the structure of the monomer investigated herein is extremely stable under ionic strengths ranging from 0 to 225 mM. The results show that a protein surface frozen owing to interactions between charged groups is mainly responsible for this stabilization. Thus, focusing on binding sites and ligands that destabilize these properties can aid the search for dengue virus inhibitors.

**Key words:** Dengue glycoprotein E; Molecular simulation; Flavivirus; Ionic strength; Unprotonated histidine residues

## INTRODUCTION

Dengue fever is a disease caused by one of the 4 dengue viruses (DENVs). The virus is transmitted by an infected mosquito within the genus *Aedes*, which occurs in tropical and subtropical areas of the world (WHO, 2012). Symptoms range from a mild fever to incapacitating high fever, and patients present with severe headache, pain behind the eyes, muscle and joint pain, and rash. No specific antiviral medicines are available for this disease. Dengue hemorrhagic fever (fever, abdominal pain, vomiting, bleeding) is a potentially lethal complication that affects mainly children. Dengue is a fast emerging, pandemic-prone viral disease in many parts of the world (WHO, 2012). It flourishes in poor urban areas, suburbs, and the countryside, but also affects residents of more affluent neighborhoods in both tropical and subtropical countries. The incidence of dengue has increased 30-fold over the last 50 years, and 50-100 million infections are now estimated to occur annually in more than 100 endemic countries, putting almost half of the world population at risk of contracting the disease (WHO, 2012). Currently, no vaccine protects against dengue, although dengue vaccines have been under development since the 1940s. The lack of vaccine is due to the limited appreciation of the global dengue disease burden and the lack of potential markets for a dengue vaccine, which has caused industry interest to languish throughout much of the 20th century. In recent years, however, dramatic acceleration has occurred in the development of a dengue vaccine (DVI, 2012).

DENV is a single positive-stranded (10.7 kb) RNA virus of the family Flaviviridae, genus Flavivirus (arbovirus type B) (Becker, 1982; Flint et al., 2004). The mature DENV consists of an external glycoprotein shell and an internal lipid bilayer (host-derived) that encloses an RNA-protein core containing the genetic material (Pokidysheva et al., 2006). The outer glycoprotein shell is composed of 180 copies of an envelope (E) glycoprotein, each of which is anchored onto the lipid bilayer through a transmembrane protein (Pokidysheva et al., 2006). The E protein is multifunctional and acts in the recognition of neutralizing antibodies and cell receptors in the morphogenesis of the virion, in the neutralization for the induction of protective immunity and the elimination of antibody responses against viral infection, and in fusion with the cell membrane (Chen et al., 1997; Germe et al., 2002; Kuhn et al., 2002; Modis et al., 2003, 2004; Zhang et al., 2004). The crystal structure of protein E of DENV type 2 (DENV2) was obtained after its expression in *Drosophila melanogaster*, and it has been described to contain 3 distinct domains - namely D-I, D-II, and D-III (Modis et al., 2003, 2005). D-I is inserted between D-II and D-III, and its primary structure is mixed with that of D-III. The relationship between D-I and D-II gives rise to a joint that provides flexibility to the E protein that seems to be important for the fusion of the membrane (Zhang et al., 2004). D-II is the dimerization domain and contains the fusion peptide. It is an elongated structure shaped like a finger, and its end consists of a sequence of hydrophobic amino acids that is conserved in all Flaviviridae. This sequence of the E protein (98-109; 98-112 in the case of DENV2) (Modis et al., 2004; Stauffer et al., 2008) is thought to be responsible for the binding of the E protein to a membrane receptor.

After virus binds to the cell surface, a complex process of viral replication is initiated by endocytosis. Once in the endosome, the surface proteins of the virus trigger membrane fusion (Stiasny and Heinz, 2006). Briefly, when the E protein dimer is exposed to low pH (around 6.5) (Modis et al., 2004), it dissociates (van der Schaar et al., 2007) and undergoes major conformational rearrangement culminating in the formation of trimers (Allison et al.,

2001). This process leads to a lump in a position that allows membrane fusion. The proposed model indicates that 3 fusion peptides penetrate the cell membrane. In the dimeric form of the E protein, the fusion peptide is accommodated in a hydrophobic region, so it is not influenced by intensive environmental conditions. D-III is a stable immunoglobulin-like fold that mediates receptor binding and concentrates a significant number of mutations, which potentially differentiate between the 4 dengue serotypes. DENV2 E protein trimers have been prepared *in vitro* using a method previously developed for protein and tick-borne encephalitis virus. In this method, homodimers are formed by soluble proteins and dissociated (reversibly) in acidic pH. The fusion cycle at low pH has been demonstrated, and the binding of the protein monomer to the lipid membrane has been shown to catalyze the irreversible formation of trimers (Modis et al., 2004). The protein reorganizes during the formation of the trimer, so that the fusion peptides merge as a single hydrophobic finger opposite D-I and D-III. The complete fusion proceeds in several steps, including dissociation of the dimer, which is followed by a trimerization process that finally brings together the monomers in the proper position for effective integration into the host cell membrane providing suitable conditions for the fusion between the virus and the membrane cavity (Stauffer et al., 2008).

The great interest in the identification and characterization of protein structure from a dynamic viewpoint is justified by the fundamental principle of molecular biology, which states that protein structure and function are closely related (Flint et al., 2004; Modis et al., 2005; Volk et al. 2007; Zhang et al., 2007). A full understanding of biological reactions is impossible without knowing exactly which molecules are involved and their roles. The relevance of studies on the identification and understanding of the structures and therefore the functions of proteins is explained by their wide application (Leah, 2001; Navarro-Sanchez et al., 2003), mainly in the chemical and pharmaceutical industries, as well as their obvious contribution to the interpretation of phenomena occurring in living matter. The development of new drugs, pesticides, herbicides, and enzymes, and the use of findings through genetic sequencing are known examples that require full knowledge and understanding of protein structure and function (Leah, 2001; Li et al., 2008; Perera et al., 2008).

At some point in the process of dengue infection, at least a fraction of the E protein is in its monomeric form - more specifically, in acidic medium and under ionic strength (IS) of approximately 150 mM (IS150). In this study, the characterization of the monomer of the DENV2 E protein before its insertion into acidic medium was undertaken. To this end, the monomer of the E protein with all the histidine residues in the unprotonated state,  $MPE_{ump}$ , was treated. In fact, the basic strategy for the molecular simulation of dengue adopted herein focused on a specific protein that plays a key role in the fusion of the virus with the host membrane so that ligands that inhibit the corresponding activity can be identified. Knowledge about the three-dimensional structure of the protein at the atomic level is essential to achieving this aim, and the E protein meets this requirement. In a previous study (Degrève et al., 2011b), the monomer of protein E of DENV2 with complete histidine residues protonation - namely,  $MPE_{prot}$  - was explored using molecular simulation.  $MPE_{prot}$  had completely distinct behavior depending on the IS of the medium in which it was examined. IS150, which is the closest to the physiological state, was the system that furnished the greatest difference in the structures obtained during the  $MPE_{prot}$  simulations and also in relation to the experimental structure. The conformation of  $MPE_{prot}$  under IS150 conditions was different mainly in terms of the angle defined by the centers of domains I and III and the axis of domain II, which was significantly

reduced. This structural change is essential for the formation of the trimer in its active form and its insertion into the cell membrane (Modis et al., 2004). The present study also examined the structural and dynamic behavior of the E protein under various IS conditions but mimicking, by the  $MPE_{\text{unp}}$  model, the extracellular neutral pH environment condition prior to fusion.

## MATERIAL AND METHODS

The study of the influence of IS on the properties of the E protein monomer of DENV was developed from data produced by simulations of the behavior of this protein in aqueous media of various IS. The interactions between atoms have been described by the GROMOS96 43a1 force field (van Gunsteren et al., 1996). The molecular dynamics simulations were performed using Gromacs 4.5.3 (Hess et al., 2008; van der Spoel et al., 2010) keeping constant the number of particles, the pressure, and the temperature (NpT ensemble) at a temperature of 300 K and a pressure of 1.0 atm using the Berendsen algorithms for temperature and pressure control (Berendsen et al., 1984). A cutoff on van der Waals interactions was applied at 1.2 nm, and the electrostatic interactions were calculated by means of the smooth particle mesh Ewald method (Essman et al., 1995). The lengths of the bonds formed with hydrogen atoms were controlled with the LINCS algorithm (Hess et al., 1997), and the structure of the water molecules was controlled with the SETTLE algorithm (Miyamoto and Kollman, 1992). The experimental conditions were adjusted according to the consideration that all the histidine residues were unprotonated. The extracellular medium under physiological conditions is approximately IS150, which is maintained mainly by sodium cations and chloride anions (Alberts et al., 2002). IS0, IS75, IS150, and IS225 were defined by appropriate concentrations of sodium and chloride ions. These values were used to understand the influence of the IS of the medium on the configuration of the protein by mimicking the extracellular conditions.

One of the monomers of the experimental structure of the DENV2 dimer obtained with X-ray diffraction (Modis et al., 2003) deposited in the Protein Data Bank (Berman et al., 2000; PDB, 2012) under the ID 1OKE was used as an initial structure and is referred to herein as Xtal. The location of the hydrogen atoms in the side groups of the histidine residues was obtained using the MolProbity program (Chen et al., 2010). The atoms ND1 of histidine residues 27 and 317 and NE2 in the other histidine residues were protonated to maintain the electroneutrality of the histidine residue side groups. After a period of relaxation, the simulations of the 4 systems at each IS were continued for 400 ns.

## RESULTS AND DISCUSSION

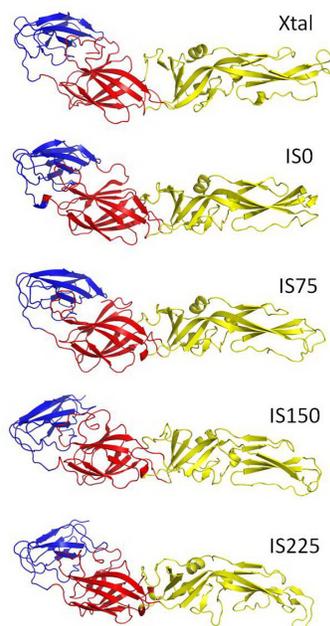
The composition of the systems and the overall main control results are listed in Table 1. The system also contained the protein monomer composed of 394 residues and 3866 atoms in the united atom force field GROMOS96 43a1. The compositions of the systems were set so that electroneutral systems would be achieved under the predefined IS. The average temperature and pressure were in accordance with the imposed conditions. The potential energies varied with the composition of the systems, but they decreased regularly with increasing IS. Indeed, the average energies by volume unit revealed a reduction of about  $9 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{nm}^{-3}$  for each 75 mM rise in IS.

**Table 1.** Composition of the systems that contain one MPE<sub>unp</sub> monomer molecule, water molecules (NH<sub>2</sub>O), sodium cations (NNa<sup>+</sup>), and chloride anions (NCl<sup>-</sup>).

	NH <sub>2</sub> O	NNa <sup>+</sup>	NCl <sup>-</sup>	$\langle E_p \rangle$	$\langle T \rangle$	$\langle p \rangle$	$\langle V \rangle$	$\langle E_p \rangle / \langle V \rangle$
IS0	36462	1	0	-0.319	300.003	1.001	1064.3	-299.9
IS75	36358	53	52	-0.347	300.005	0.998	1123.4	-308.9
IS150	36254	105	104	-0.353	300.008	1.001	1110.7	-317.5
IS225	33831	157	156	-0.353	300.009	1.001	1079.9	-326.7

$\langle E_p \rangle$  = potential energy (10<sup>6</sup> kcal/mol);  $\langle T \rangle$  = temperature (K);  $\langle p \rangle$  = pressure (atm);  $\langle V \rangle$  = volume (nm<sup>3</sup>), and  $\langle E_p \rangle / \langle V \rangle$  = potential energy by volume unit (kcal·mol<sup>-1</sup>·nm<sup>-3</sup>).

Figure 1 depicts the average time structures of the monomer, which corresponded to the ensemble mean structures (EMS) as well as to Xtal. As shown, the structure in IS225 is similar to, but a little more compact, than that of Xtal. This difference can be explained by the fact that the medium of IS225 is the closest to the experimental medium, bearing in mind that the X-ray diffraction data were acquired from crystals produced in a basic medium of high IS (Modis et al., 2003). D-I and D-III were similar in all situations, but D-II was rotated under IS0, IS75, and IS150 compared with the positioning of Xtal and IS225. Moreover, D-II became more compact as the IS increased. These changes are reflected in the root-mean-square deviations (RMSD) listed in Table 2. The average simulation structures exponentially deviated from that of Xtal as a function of the IS:  $\text{RMSD (nm)} = 0.33 \pm 0.06 \exp(x/118)$ , where x is the IS in millimoles, with an adjusted R<sup>2</sup> of 0.995. The RMSD between the simulated structures also increased as a function of the difference in the respective IS. Apparently, the IS did not promote structural changes under any circumstances, as reflected by the RMSD values, which did not evolve gradually.



**Figure 1.** Xtal and EMS structures produced by the simulations. All the EMS structures were superimposed with the Xtal experimental structure. Domains I, II, and III are identified by red, yellow, and blue colors, respectively. The figures were produced with the PyMol program (PyMol, 2012).

**Table 2.** Root-mean-square deviation between the various ensemble mean structure and the Xtal structures (nm).

	Xtal	IS0	IS75	IS150	IS225
Xtal	0	0.39	0.46	0.55	0.75
IS0	0.39	0	0.34	0.53	0.73
IS75	0.46	0.34	0	0.52	0.67
IS150	0.55	0.53	0.52	0	0.54
IS225	0.75	0.73	0.67	0.54	0

The same conclusions can be drawn from Table 3, which shows that the gyration radii are virtually independent of the IS. However, compared with the volume of Xtal, the volume of the spheres displaying radii equal to the gyration radii showed an 11% compression. Conversely, the largest distances between the protein atoms ( $D_{\max}$ ) showed no clear differences, which indicated that the dissimilarity in the structures must be located in the intermediate, or internal, regions. The rotations observed for D-II must consequently occur in intermediate regions and do not affect the overall protein structure.

**Table 3.** Average radii of gyration,  $\langle R_g \rangle$ , and the largest distance (D) between the atoms of the monomer of protein E.

	$\langle R_g \rangle$ (nm)	$D_{\max}$ (nm)
Xtal	3.42	12.0
IS0	3.29	11.6
IS75	3.30	11.8
IS150	3.27	12.1
IS225	3.28	11.8

## Secondary structures

Secondary structures were defined based on the identification of hydrogen bonds between atoms in the main chain (Kabsch and Sander, 1983) under IS0, IS75, IS150, and IS225 EMS as well as in Xtal. The results are summarized in Table 4. All of the experimentally established secondary structures were also found in the simulated systems, but the antiparallel sheet (374-380) (387-393) of the C-terminal region had no experimental correspondence. Several differences among the IS structures were found, such as a short sheet in D-I of IS0 and in the fusion peptide in IS225 D-II, as well as loss of sheets - 2 in IS150, 1 in D-I, 1 in D-II, and 1 in IS225 D-II.

Some residues participated in more than one secondary structure, thereby increasing the stability of the particular region. This kind of formation has been called the extended secondary structure (ESS) (Degrève et al., 2011a). The ESS listed in Table 5 for  $MPE_{\text{unp}}$  were shorter than those found in  $MPE_{\text{prot}}$  (Degrève et al., 2011a), showing that the configuration of the monomer of the E protein is much less stable or less locked when the protonation of the histidine residues is incomplete. This characteristic is one more indication of the role of these residues in the process of virus fusion with the cell membrane. The ESS are very stable, as shown in Figure 2, in which the distributions of the specific RMSD of the various ESS are represented - namely  $D_{\text{ESS}}$  (RMSD). These RMSD were calculated using EMS as references. The mean values of the RMSD of the ESS calculated along 400 ns are also listed in Table 5. Their small values (less

than 0.1 nm) confirmed that the regions identified by the ESS were highly stable and should therefore have an important role in the overall stability of the molecule.

**Table 4.** Residues that form the secondary structures in the ensemble mean structure and in Xtal.

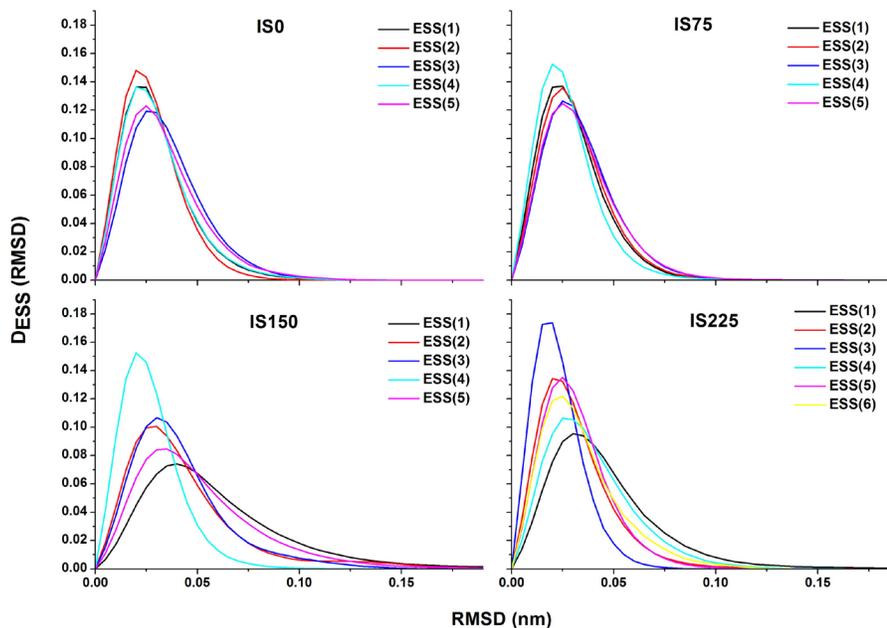
Xtal	ISO	IS75	IS150	IS225
Antiparallel sheets				
<b>20-25</b>	<b>283-288</b>	<b>21-26</b>	<b>282-287</b>	<b>21-25</b>
31-33	41-43	29-33	41-45	31-33
<b>41-45</b>	<b>139-143</b>	<b>40-44</b>	<b>140-144</b>	<b>40-45</b>
47-50	135-138	47-49	135-137	139-144
		48-50	276-278	40-45
<b>57-59</b>	<b>220-222</b>	<b>57-59</b>	<b>220-222</b>	<b>57-59</b>
<b>62-66</b>	<b>118-123</b>	<b>62-65</b>	<b>119-123</b>	<b>62-65</b>
69-72	113-116	68-73	112-117	119-123
<b>90-99</b>	<b>109-118</b>	<b>90-99</b>	<b>109-118</b>	<b>90-99</b>
				109-118
				97-99
<b>126-128</b>	<b>198-200</b>	<b>126-129</b>	<b>197-200</b>	<b>126-129</b>
138-141	160-163	138-141	160-163	138-141
<b>171-175</b>	<b>179-183</b>	<b>170-175</b>	<b>179-184</b>	<b>171-175</b>
<b>182-186</b>	<b>284-288</b>	<b>182-187</b>	<b>283-288</b>	<b>182-187</b>
<b>196-200</b>	<b>205-209</b>	<b>196-200</b>	<b>205-209</b>	<b>196-200</b>
205-207	268-271	205-207	268-271	205-209
<b>238-241</b>	<b>249-252</b>	<b>238-241</b>	<b>249-252</b>	<b>238-241</b>
<b>306-308</b>	<b>324-326</b>	<b>305-308</b>	<b>324-327</b>	<b>306-308</b>
<b>320-324</b>	<b>365-369</b>	<b>320-324</b>	<b>365-369</b>	<b>320-324</b>
<b>337-339</b>	<b>378-380</b>	<b>337-339</b>	<b>378-380</b>	<b>337-339</b>
		374-380	387-393	378-380
				375-380
				387-392
Parallel sheets				
<b>9-13</b>	<b>30-34</b>	<b>8-13</b>	<b>29-34</b>	<b>8-14</b>
				<b>29-35</b>
				<b>9-14</b>
				<b>30-35</b>
				<b>11-13</b>
				<b>32-34</b>
Helix				
	<b>256-260</b>	<b>256-260</b>	<b>256-260</b>	<b>256-260</b>
				<b>257-262</b>

Note that only one helix is found among the criteria used. The secondary structures recurrent in all systems are highlighted in bold.

**Table 5.**  $MPE_{\text{unp}}$  extended secondary structure (ESS) in the four systems.

	ISO	IS75	IS150	IS225
ESS(1)	8-13, 29-34, 41-45 (17, 3.3 ± 1.7)	21-25, 181-187, 283-289 (19, 3.2 ± 1.6)	9-13, 30-34, 40-48, 137-144, 158-163 (33, 6.4 ± 4.1)	40-50, 135-144, 276, 277 (23, 4.8 ± 3.0)
ESS(2)	126-129, 196-200, 205-209, 268, 269 (16, 3.0 ± 1.5)	126-129, 196-209, 268, 269 (20, 3.4 ± 1.6)	21-25, 182-187, 283-288 (17, 4.7 ± 3.1)	68-73, 90-93, 112-118 (17, 3.3 ± 1.8)
ESS(3)	170-175, 179-187, 283-288 (21, 3.8 ± 1.9)	171-175, 179-183, 291, 292 (12, 3.6 ± 1.7)	125-129, 196-201, 205-209 (16, 4.4 ± 2.4)	126-129, 196-200, 205-209 (14, 2.5 ± 1.2)
ESS(4)	305-308, 320-327, 365-369 (17, 3.3 ± 1.7)	305-308, 320-327, 365-369 (17, 3.0 ± 1.5)	305-308, 320-327, 365-369 (17, 3.0 ± 1.5)	171-175, 179-185, 285-288 (16, 4.1 ± 2.2)
ESS(5)	337-339, 374-380, 387-390 (14, 3.6 ± 1.9)	337-339, 374-380, 387-393 (17, 3.6 ± 1.8)	337-339, 374-381, 386-393 (19, 5.3 ± 3.0)	306-308, 320-326, 365-369 (15, 3.4 ± 1.7)
ESS(6)				337-339, 377-380, 387-390 (11, 3.7 ± 2.0)

The total number of residues for each ESS and the average root-mean-square deviation (referent to ensemble mean structure) and corresponding standard deviation for each ESS (units of 0.01 nm) are shown in parentheses.



**Figure 2.** Distributions,  $D_{ESS}$  (RMSD), of the specific RMSD for the various extended secondary structure (ESS) using the ensemble mean structure (EMS) as references. RMSD = root-mean-square deviation.

### Hydrogen bonds and interaction energies

The total mean interaction energies of  $MPE_{unp}$  with the water molecules,  $E(\text{res}, H_2O)$ , were  $-10.901$ ,  $-10.391$ ,  $-10.554$ , and  $-10.541 \times 10^3$  kcal/mol under IS0, IS75, IS150, and IS225, respectively. These results were similar under the 4 IS. The residues with charged lateral groups presented more attractive  $E(\text{res}, H_2O)$  values, with an average in kilocalories per mole of  $-48.1 \pm 22.2$  (Arg),  $-101.6 \pm 31.36$  (Asp),  $-105.5 \pm 35.5$  (Glu), and  $-82.3 \pm 32.5$  (Lys), whereas the mean value for all residues was  $-26.9 \pm 39.4$  kcal/mol. The average  $E(\text{res}, H_2O)$  per residue (Table 6) was fairly constant as a function of both the IS of the medium and of the domain. The same cannot be said for the charged residues Arg, Asp, Glu, and Lys. The total mean numbers of hydrogen bonds in the monomer were 2889.4, 2859.1, 2887.5, and 2886.7 under IS0, IS75, IS150, and IS225, respectively.

The total mean interaction energies between  $MPE_{unp}$  and the chlorine ions were  $-264.4$ ,  $-412.4$ , and  $-558.8$  kcal/mol under IS75, IS150, and IS225, respectively. The interaction energies between the residues and ions,  $E(\text{res}, \text{ion})$ , occurred throughout the chain, especially with the charged side groups, as shown in Table 7, which lists the residues that have at least one  $E(\text{res}, \text{ion})$  equal to or less than  $-10.0$  kcal/mol. Generally,  $E(\text{res}, \text{ion})$  became more attractive as IS increased. This trend was seen clearly in the number of ions forming the ionic layers of the residues: 44.7 sodium and 39.6 chlorine in IS75, and 80.8 and 77.9 ions in IS150 and 144.9 and 116.2 ions in IS225. However, exceptions occurred in the Lys122, Glu148, and Glu360 residues. Their  $E(\text{res}, \text{ion})$  values were virtually independent of IS, thereby indicating that their ionic layers are particularly stable. These 3 residues are located on the outer face of

MPE<sub>unp</sub>, so they are in direct contact with the external saline solution. None of the Lys122, Glu148, and Glu360 residues belongs to an ESS. Another exception was found in the cases of Lys58, Lys64, Glu84, Glu174, Asp192, Glu338, and Glu383, in which E(res,ion) was much more attractive under IS150. These residues face several sheets, so their access to the external medium depends directly on the neighborhood. For example, the distance between the side chain oxygen atoms of Glu338 and Glu383 in the IS150 EMS is 0.93 nm, so the 2 residues can share the same ions in their ionic layers. Lys58, Lys64, Glu84, Glu174, Asp192, Glu338, and Glu383 do not belong to an ESS with the exception of Glu174 and Glu338, which are included in small structures of an ESS in IS225. The opposite behavior is observed for Asp154, Glu161, Lys202, and Asp215. These residues are located in a region in which the main chain is slightly curved under IS150, but are located at the ends of almost helix-like regions under IS225 (Lys157-Thr155 and Gln211-Asp215), which increases their exposure to the external environment. Again, Asp154, Glu161, Lys202, Asp215 do not belong to an ESS.

**Table 6.** E(res,H<sub>2</sub>O) averages of residues with the charged side group and the average for all residues depending on the domain and on the ionic strength (kcal·mol<sup>-1</sup>·residue<sup>-1</sup>).

	IS0	IS75	IS150	IS25
Arg				
D-I	-47.3 (11.3)	-48.5 (10.3)	-52.6 (11.4)	-43.5 (10.9)
D-II	-45.5 (10.1)	-47.4 (9.6)	-58.5 (11.2)	-59.5 (11.1)
D-III	-36.8 (11.2)	-40.3 (11.3)	-40.5 (12.1)	-45.1 (12.5)
Asp				
D-I	-96.3 (14.6)	-88.8 (14.3)	-92.3 (15.9)	-87.6 (14.9)
D-II	-105.5 (18.7)	-104.0 (18.4)	-107.0 (19.1)	-105.3 (19.2)
D-III	-116.7 (19.0)	-105.6 (17.9)	-101.7 (17.4)	-121.1 (19.8)
Glu				
D-I	-113.1 (16.9)	-103.3 (16.4)	-101.0 (16.6)	-104.1 (16.6)
D-II	-101.5 (17.2)	-98.9 (17.1)	-104.3 (17.7)	-101.0 (17.5)
D-III	-110.8 (17.5)	-110.9 (17.2)	-111.8 (17.6)	-107.8 (17.0)
Lys				
D-I	-74.2 (9.6)	-67.2 (8.9)	-68.9 (9.1)	-73.6 (9.5)
D-II	-89.7 (9.4)	-86.4 (9.4)	-82.0 (9.2)	-78.9 (8.9)
D-III	-96.7 (11.4)	-90.3 (11.0)	-89.6 (11.6)	-90.4 (11.6)
All residues				
D-I	-28.8 (7.2)	-26.0 (7.0)	-26.3 (7.0)	-26.6 (7.1)
D-II	-25.5 (7.6)	-28.1 (7.5)	-25.6 (7.6)	-24.9 (7.4)
D-III	-29.9 (7.1)	-29.1 (7.1)	-29.4 (7.2)	-30.2 (7.5)

The average number of hydrogen bonds between the residue and the solvent are in parentheses.

**Table 7.** E(res,ion), in kcal/mol, for the residues presenting at least one E(res,ion) ≤ -10.0 kcal/mol.

	IS75	IS150	IS225		IS75	IS150	IS225		IS75	IS150	IS225
Asp10	-7.7	-8.2	-11.7	Lys123	-6.8	-11.3	-12.4	Lys247	-8.6	-10.5	-15.4
Lys36	-8.9	-8.6	-14.2	Glu148	-12.3	-13.6	-13.1	Lys310	-3.6	-5.6	-12.0
Lys58	-6.2	-16.6	-8.5	Asp154	-5.1	-1.5	-11.4	Glu314	-0.1	-5.0	-17.5
Lys64	-17.5	-32.2	-19.4	Glu161	-7.2	-2.1	-10.1	Glu338	-3.9	-10.6	-5.8
Glu71	-1.0	+1.4	-13.6	Glu172	-4.2	-6.6	-13.9	Asp341	-1.5	-0.9	-13.1
Glu84	-1.0	-11.2	-6.0	Glu174	-4.6	-11.4	-9.6	Glu360	-18.7	-18.9	-19.5
Glu85	-4.4	-6.4	-14.6	Asp192	-2.8	-10.6	-9.4	Asp362	-11.3	-12.9	-3.9
Arg89	-5.5	-8.4	-11.0	Lys202	-15.3	-11.6	-22.4	Glu370	-0.7	-2.4	-12.9
Arg99	-2.8	-8.8	-10.9	Asp215	-4.5	-1.9	-10.8	Glu383	-6.2	-10.4	-4.3
Lys122	-12.3	-11.0	-11.7	Lys246	-12.2	-12.8	-19.8	Lys394	-4.6	-7.1	-10.0

Finally, extremely attractive energies were found for the Lys64, Lys202, Glu314, and Glu360 residues. In the 1OKE structure of the dimer of the E protein, the Asp154 residue of one monomer is very close to the Lys247 residue of another monomer ( $\sim 0.86$  nm). These results indicate that a clear separation exists between the contributions of the intramolecular and intermolecular interaction energies in the definition and maintenance of the tertiary structure of  $MPE_{\text{unp}}$ .

### Stability of the structures

The state of the protein in which no histidine residue is protonated corresponds to that observed in non-acidic medium before the introduction of the virus into the target body. The outer layer of the virus should therefore be structurally very stable, which implies that the 180 copies of the E protein are too rigid. This behavior can be checked by measuring the distances between electrically charged side groups, for instance, because the electrostatic interactions are determining factors in the stabilization of structures. Table 8 depicts the most important pairs of residues with charged side groups with average distances less than 0.6 nm as well as the corresponding standard deviations. Notably, several pairs of residues are found at virtually fixed distances in all systems. The small standard deviations suggest that these pairs of residues make the surface of the protein nearly frozen. In particular, the Glu62, Glu126, Lys58, Lys122, and Lys123 residues maintain the D-II region, to which the fusion peptide binds rigidly, which facilitates the maintenance of the fusion peptide structure necessary in the pre-fusion phase. Interestingly, the protonation, or lack of protonation, of the histidine residues results in opposite behaviors: the lack of protonation decreases the ESS and increases the rigidity of the surface of the protein related to the protonated systems (Degrève et al., 2011a).

**Table 8.** The most representative average distances between charged groups and their variances (after vertical bars) in nm.

Pairs		IS0	IS75	IS150	IS225
Asp22	Lys284	0.49 0.12	0.51 0.13	0.50 0.09	0.52 0.13
Asp98	Lys110		0.55 0.16		0.42 0.08
Glu44	Arg2			0.59 0.10	0.56 0.09
Glu184	Arg286	0.55 0.11	0.54 0.12	0.58 0.11	0.52 0.09
Glu368	Arg9		0.54 0.04	0.53 0.08	0.50 0.06
Glu370	Arg350	0.52 0.12	0.56 0.17	0.51 0.13	0.51 0.14
Glu49	Lys163	0.40 0.07	0.41 0.08		0.45 0.12
Glu62	Lys122	0.50 0.11	0.52 0.14	0.51 0.07	0.53 0.11
Glu62	Lys123	0.50 0.09	0.47 0.09	0.56 0.14	0.52 0.12
Glu126	Lys58	0.47 0.08	0.48 0.07	0.46 0.08	0.52 0.08
Glu136	Lys51	0.43 0.08	0.44 0.11	0.56 0.18	0.51 0.14
Glu136	Lys163	0.55 0.09	0.52 0.10	0.44 0.08	0.50 0.09
Glu147	Lys295	0.58 0.10	0.51 0.16	0.42 0.04	
Glu269	Lys204	0.44 0.11	0.49 0.16	0.52 0.17	0.52 0.14
Glu327	Lys307	0.51 0.12	0.50 0.11	0.52 0.11	0.55 0.12
Glu370	Lys393	0.47 0.12	0.50 0.11		

The atoms used to determine the distances between groups are CZ for Arg, CG for Asp, CD for Glu, and NZ for Lys.

The present study on the behavior of  $MEP_{\text{unp}}$  shows that the protein structure is extremely stable under IS0-IS225 when all histidine residues are unprotonated. The main features of the secondary structures encountered in  $MEP_{\text{unp}}$  are also found in the experimental

structure and in the same system when all the histidine residues are protonated. However, the ESS structures in MEP<sub>unp</sub> are smaller than those of MEP<sub>prot</sub> proteins. Nevertheless, the MEP<sub>unp</sub> surface is largely frozen by interactions between charged residues on the protein surface. These interactions can occur directly or indirectly through the ionic layers of the charged groups and, of course, through both hydration layers of the ions and of the charged groups. Comparison of the data obtained for MEP<sub>unp</sub> with those for MEP<sub>prot</sub> showed that the protein was extremely rigid in the first protonation scheme and that this rigidity was much more a consequence of the interactions between the surface groups of the protein and between these groups and the external environment than of interactions between internal regions. The stabilizing contributions in MEP<sub>prot</sub> are opposite. Because the external medium - namely solvent and other components - are key factors in determining and maintaining the structure of proteins, MEP<sub>prot</sub> can be expected to be less rigid than MEP<sub>unp</sub> is, which is observed because the structure of MEP<sub>prot</sub> changes under IS150 - that is, a fold that is essential for the formation of the trimer in the pre-fusion phase. Conversely, the rigidity of MEP<sub>unp</sub> is necessary to avoid the destruction of the structures of the virus E protein. If charged groups are partially responsible for the rigidity of the MEP<sub>unp</sub> structure, they could hardly be targets for specific inhibitors, because electrostatic interactions are nonspecific. However, because pairs of charges are involved, they should be considered when seeking ligands that could destabilize the structures of protein E in the MEP<sub>unp</sub> form, aiming to destroy the structure of the DENV E protein itself and inhibiting it even in the mature virus.

## ACKNOWLEDGMENTS

Research supported by Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico.

## REFERENCES

- Alberts B, Johnson A, Lewis J, Raff M, et al. (2002). *Molecular Biology of the Cell*. Garland Science, New York.
- Allison SL, Schlich J, Stiasny K, Mandl CW, et al. (2001). Mutational evidence for an internal fusion peptide in flavivirus envelope protein E. *J. Virol.* 75: 4268-4275.
- Becker Y (1982). *Molecular Virology: Molecular and Medical Aspects of Disease-Causing Viruses of Man and Animals*. Martinus Nijhoff Publishers, Hingham.
- Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, et al. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81: 3684-3690.
- Berman HM, Westbrook J, Feng Z, Gilliland G, et al. (2000). The Protein Data Bank. *Nucleic Acids Res.* 28: 235-242.
- Chen VB, Arendall WB, III, Headd JJ, Keedy DA, et al. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66: 12-21.
- Chen Y, Maguire T, Hileman RE, Fromm JR, et al. (1997). Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nat. Med.* 3: 866-871.
- Degrève L, Fuzo CA, Silva GM and Caliri A (2011a). The Stability of Extended Secondary Protein Structures: The Dengue Virus E Protein Case. In: Ninth Triennial Congress of the World Association of Theoretical and Computational Chemists Santiago de Compostela, PII-083.
- Degrève L, Fuzo CA and Caliri A (2011b). Molecular simulation of the e protein from dengue virus type 2. 36th FEBS Congress of the Biochemistry for Tomorrows Medicine. *FEBS J.* 278: 368-368.
- DVI (2012). Dengue Vaccine Initiative. Available at [<http://www.denguevaccines.org/why-a-vaccine>]. Accessed February 6, 2012.
- Essman U, Perera L, Berkowitz ML, Darden T, et al. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.* 103: 8577-8592.

- Flint SJ, Enquist LW, Racaniello VRT and Skalka AM (2004). Principles of Virology: Molecular Biology, Pathogenesis, and Control of Animal Viruses. ASM Press, Washington.
- Germi R, Crance JM, Garin D, Guimet J, et al. (2002). Heparan sulfate-mediated binding of infectious dengue virus type 2 and yellow fever virus. *Virology* 292: 162-168.
- Hess B, Bekker H, Berendsen HJC and Fraaije JGEM (1997). LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 18: 1463-1472.
- Hess B, Kutzner C, van der Spoel D and Lindahl E (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4: 435-447.
- Kabsch W and Sander C (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22: 2577-2637.
- Kuhn RJ, Zhang W, Rossmann MG, Pletnev SV, et al. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell* 108: 717-725.
- Leah AR (2001). Molecular Modeling: Principles and Applications. Prentice Hall, New York.
- Li Z, Khaliq M, Zhou Z, Post CB, et al. (2008). Design, synthesis, and biological evaluation of antiviral agents targeting flavivirus envelope proteins. *J. Med. Chem.* 51: 4660-4671.
- Miyamoto S and Kollman PA (1992). SETTLE - an analytical version of the shake and rattle algorithm for rigid water models. *J. Comput. Chem.* 13: 952-962.
- Modis Y, Ogata S, Clements D and Harrison SC (2003). A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* 100: 6986-6991.
- Modis Y, Ogata S, Clements D and Harrison SC (2004). Structure of the dengue virus envelope protein after membrane fusion. *Nature* 427: 313-319.
- Modis Y, Ogata S, Clements D and Harrison SC (2005). Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. *J. Virol.* 79: 1223-1231.
- Navarro-Sanchez E, Altmeyer R, Amara A, Schwartz O, et al. (2003). Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses. *EMBO Rep.* 4: 723-728.
- PDB (2012). Protein Data Bank. Available at [<http://www.pdb.org>]. Accessed February 6, 2012.
- Perera R, Khaliq M and Kuhn RJ (2008). Closing the door on flaviviruses: entry as a target for antiviral drug design. *Antiviral Res.* 80: 11-22.
- Pokidysheva E, Zhang Y, Battisti AJ, Bator-Kelly CM, et al. (2006). Cryo-EM reconstruction of dengue virus in complex with the carbohydrate recognition domain of DC-SIGN. *Cell* 124: 485-493.
- PyMol (2012). The PyMOL Molecular Graphics System. Available at [<http://sourceforge.net/projects/pymol/files/Legacy/>]. Accessed February 6, 2012.
- Stauffer F, Melo MN, Carneiro FA, Sousa FJ, et al. (2008). Interaction between dengue virus fusion peptide and lipid bilayers depends on peptide clustering. *Mol. Membr. Biol.* 25: 128-138.
- Stiasny K and Heinz FX (2006). Flavivirus membrane fusion. *J. Gen. Virol.* 87: 2755-2766.
- van der Schaar HM, Rust MJ, Waarts BL, Ende-Metselaar H, et al. (2007). Characterization of the early events in dengue virus cell entry by biochemical assays and single-virus tracking. *J. Virol.* 81: 12019-12028.
- van der Spoel D, Lindahl E, Hess B, van Buuren AR, et al. (2010). Gromacs User Manual Version 4.5.4. Available at [<http://www.gromacs.org>]. Accessed February 6, 2012.
- van Gunsteren WF, Krüger P, Billeter SR, Mark AE, et al. (1996). Biomolecular Simulation: The Gromos 96 Manual and User Guide. Hoehchulverlag an du ETH, Zürich.
- Volk DE, Lee YC, Li X, Thivyanathan V, et al. (2007). Solution structure of the envelope protein domain III of dengue-4 virus. *Virology* 364: 147-154.
- WHO (2012). World Health Organization. Available at [<http://www.who.int/topics/dengue/en/>]. Accessed February 6, 2012.
- Zhang J, Wang J, Gao N, Chen Z, et al. (2007). Up-regulated expression of  $\beta$ 3 integrin induced by dengue virus serotype 2 infection associated with virus entry into human dermal microvascular endothelial cells. *Biochem. Biophys. Res. Commun.* 356: 763-768.
- Zhang Y, Zhang W, Ogata S, Clements D, et al. (2004). Conformational changes of the flavivirus E glycoprotein. *Structure* 12: 1607-1618.