Stereochemical features of the envelope protein Domain III of dengue virus reveals putative antigenic site in the five-fold symmetry axis

R.O.S. Soares *, A. Caliri

Departamento de Física e Química, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Av. do Café, S/N. 14040-903, Ribeirão Preto, São Paulo, Brazil

A R T I C L E   I N F O
Article history:
Received 20 June 2012
Received in revised form 7 September 2012
Accepted 11 September 2012
Available online 23 September 2012

Keywords:
Dengue virus
Domain III
Molecular dynamics
Putative epitope
Surface pore
Induced fit

A B S T R A C T
We bring to attention a characteristic parasitic pattern present in the dengue virus: it undergoes several intensive thermodynamic variations due to host environmental changes, from a vector’s digestive tract, through the human bloodstream and intracellular medium. Comparatively, among the known dengue serotypes, we evaluate the effects that these medium variations may induce to the overall structural characteristics of the Domain III of the envelope (E) protein, checking for stereochemical congruences that could lead to the identification of immunologic relevant regions. We used molecular dynamics and principal component analysis to study the protein in solution, for all four dengue serotypes, under distinct pH and temperature.

1. Introduction

Dengue fever and its complications – dengue hemorrhagic fever and dengue shock syndrome – are currently some of the most incident mosquito-borne diseases in the world. Despite its tropical nature, spreading to more temperate habitats seems to be a tendency, as the anthropogenic-induced global climate change takes place on a global scale [1,2]. Basically two major strategies are currently employed to counter its spread: one focuses on the vectors (Aedes aegypti and Aedes albopictus) as a target of preventive population control, and the other on the virus itself, with efforts to develop effective inhibition of infection, since nowadays dengue can only be treated palliatively. The later comprises biochemical approaches to find ways to invalidate the virus’ mechanisms of infection and prevent the development of the symptoms. In spite of much progress in the past few years [3–6], there are to date no commercially viable vaccines.

Dengue virus belongs to the Flavivirus genus, which includes several other notable viruses, such as those of yellow fever, tick borne encephalitis and Japanese encephalitis. It presents a smooth icosahedron outer layer with a diameter of approximately 500 Å, composed of 180 envelope (E) glycoproteins in a 90 dimer setup [7]. Along with the E protein (a class II fusion protein [7–9]), two other structural proteins are present and occur in stoichiometric amounts; they are found in the membrane (M) and core (C) of the virus [9]. The innermost portion of C proteins form a closed surface that involves the plus-sense RNA of about ~10,700 nucleotides [7]. Here we label the four dengue serotypes as DENV-1, 2, 3 and 4.

The fusion of the viral and host membranes takes place inside a cell’s endosome, and according to the current proposed mechanism [10], it is of central importance for delivering of the viral genetic material to the cell’s replication machinery, and thus for the infection to succeed. Studies of these fusion mechanisms have indicated a key role of the histidine residues in detecting the pH lowering inside the endosome [11–13]. In such acid environment, the envelope proteins suffer major structural changes leading to their trimerization, leading to formation of membrane-fusion structures on the viral surface [7,10,14]. Of the five preserved histidines among the flaviviruses, one (His323) is found at the interface between Domain I (DI) and Domain II (DII) of protein E [12], which is allegedly determinant to the dynamics of the Domain III (DIII) during viral/cell membrane fusion, because of its observed displacement and rotation relative to Domain I [15,16].

The object of our study, DIII, has been demonstrated as the target for antibodies [17–20] and accumulates a large number of mutations...
throughout the four dengue serotypes. Moreover, this domain is remarkable for its primary structure contiguousness (a characteristic not intrinsic to the other two domains), as well as its stand-alone stability in the absence of its neighbor domains, DI and DII.

For several years now, in silico methods are increasingly employed as effective tools for approaching biochemical themes, such as protein folding, drug discovery, and binding site search [21–24], among others. In the same way, molecular dynamics has become increasingly common to evaluate the structural aspects of the dengue virus, as well as other viruses [25–28].

In this work, we present the results of the use of molecular dynamics (MD) simulations with an explicit solvent to investigate the structural behavior of the DIII in solution, through variations of intensive thermodynamic conditions, namely pH and temperature. We were able to map the structural flexibilities of the DIII chain, and associate new-found particularities with the current literature to propose immunological-relevant insights.

2. Material and methods

2.1. System

The protein coordinates where obtained from the RCSB Protein Data Bank [29] | DENV-1 (PDB ID: 3G7T), DENV-2 (PDB ID: 2SF), DENV-3 (PDB ID: 1UZG) and DENV-4 (PDB ID: 2H0P) | [30–33]. The structures of DENV-2 and 4 were determined from isolated DIII by nuclear magnetic resonance (NMR). For serotypes 1 and 3, however, the coordinates of DIII were extracted from the structures of the respective E proteins determined by X-ray diffraction. Therefore, appropriate adaption for a proper simulation system was necessary, which consisted of the removal of eventual experimental artifacts (such as a poly-histidine tail of DENV-2), resulting in the coordinates from Met289-Ser396 for DENV-1/DIII and Met287-Lys392 for DENV-3/DIII. Aiming to preserve the integrity of the experimental structures, we maintained the (nine) residues from the DI/DIII hinge that were present in DENV-2 and DENV-4. Moreover, these correspondent residues were also maintained when extracting the DIII from serotypes 1 and 3. After this treatment, the fragment was submitted to energy minimization, causing the aforementioned hinge that links domains I and III (N-terminus region of the isolated DIII) to adjust itself to a more energetically favorable conformation. As a complementary analysis we set a system which consisted of five DIIIIs aligned in a symmetric 5-fold disposition, much alike to the one found on the virus surface. Initially, we obtained the coordinates from an assembly of E proteins at the same ordered setup such as the one found in the virus envelope (PDB ID: 1THD [14], DENV-2). We then used this assembly’s spatial location and angles among the monomers, to fit multiple copies of a higher resolution structure for the same serotype (PDB ID: 1OAN [34]), thus obtaining a hybrid system with a proper spatial setting and atomic resolution. Afterwards, we isolated only the five DIIIIs and, since the subunits are now loose, we proceed with the restraining of movements of a few backbone atoms of each domain (from Val309 to Cys333 and Val354 to Pro371), those which are farthest from the center of the pore (icosahedron vertex). This restrain is placed in a manner so that all the atoms of interest (inner loop regions) are least affected, and the quaternary structure maintains its original setting as found on the virus surface.

2.2. Simulation parameters

In the present work, the simulation systems comprehend the solvated DIII of the E protein for each of the four DENV serotypes, submitted to diverse intensive conditions. The temperature was considered at two levels: 298 K, which mimics the environment temperature, and thus the same as the digestive tract of the mosquito (an ectothermic animal); and 310 K, which mimics the conditions in the human non-peripheral blood stream and internal organs. Thus the evaluation of DIII stereochemical behavior as a function of the temperature, from environmental to physiological, could reveal features that may occur during the process of infection. Similarly, we set three pH values for each temperature, corresponding to the environmental acidification that the virus undergoes from its neutral vector intestinal tract and salivary glands through the human bloodstream and acidic endosome. For practical methodological terms, the pH values were set to 3, 5 and 7 by fixing the protonation states of ionizable side chains in its correspondent pka. Since we have four serotypes with three variations on pH and two on the temperature, this section comprises 24 simulations, with 200 ns duration each.

For all simulations and analyses we used the molecular modeling package GROMACS [35] (v. 4.5.x) in a Silicon Graphics International® Altix® XE 1300 cluster. The all-atom molecular dynamics was carried out under the OPLS-AA force field [36] and the PME (particle mesh Ewald) was used to calculate electrostatic interactions. The employed explicit solvent was the SPC [37] water model, with a subset of control simulations (not reported here) where the TIP4P [38] water model was utilized and converged to very similar results. Protein coordinates were set in cubic boxes with the sides of length around 7.5 nm, in such a manner that, at the initial position, the protein was centralized and the extremities of its longer axis were 1.2 nm far from the cubic faces. Each system consisted a distinct amount of atoms, as represented in Table 1. Ions were provided primarily in a proper amount in order to keep a net charge equal to zero, according to the ratio of protein ionized groups for each corresponding pH. Since we treated the DIII in equilibrium (and no large chain deformations were expected), the pH for each system was set indirectly, ionizing the residues based on their side-chain pK values (according to [39]) at the beginning of the simulations.

First, we ran energy minimization simulations of the systems, through the steepest descent method, including the SETTLE algorithm [40], until the converging of a threshold smaller than 2 × 10⁻³ kJ mol⁻¹ nm⁻¹. Next, we used the Maxwell–Boltzmann distribution to generate the initial

| DENV-1 | pH 3 | 41,474 | 1693 | 39,765 | 0 | 16 | 13K, 1R, 1H, 1T₉ | – |
| DENV-1 | pH 5 | 41,429 | 1678 | 39,720 | 15 | 16 | 13K, 1R, 1H, 1T₉ | 4D, 10E, 1T₉ |
| DENV-1 | pH 7 | 41,433 | 1677 | 39,726 | 15 | 15 | 13K, 1R, 1T₉ | 4D, 10E, 1T₉ |
| DENV-2 | pH 3 | 38,949 | 1732 | 37,200 | 0 | 17 | 11K, 3R, 2H, 1T₉ | – |
| DENV-2 | pH 5 | 38,514 | 1717 | 36,765 | 15 | 17 | 11K, 3R, 2H, 1T₉ | 5D, 9E, 1T₉ |
| DENV-2 | pH 7 | 38,507 | 1715 | 36,762 | 15 | 15 | 11K, 3R, 1T₀ | 5D, 9E, 1T₉ |
| DENV-3 | pH 3 | 38,963 | 1678 | 37,266 | 0 | 19 | 14K, 2R, 2H, 1T₉ | – |
| DENV-3 | pH 5 | 38,921 | 1663 | 37,224 | 15 | 19 | 14K, 2R, 2H, 1T₉ | 4D, 10E, 1T₉ |
| DENV-3 | pH 7 | 38,920 | 1661 | 37,227 | 15 | 17 | 14K, 2R, 1T₀ | 4D, 10E, 1T₉ |
| DENV-4 | pH 3 | 41,494 | 1747 | 39,729 | 0 | 18 | 11K, 4R, 2H, 1T₉ | – |
| DENV-4 | pH 5 | 41,455 | 1734 | 39,690 | 13 | 18 | 11K, 4R, 2H, 1T₉ | 3D, 9E, 1T₉ |
| DENV-4 | pH 7 | 41,460 | 1732 | 39,699 | 13 | 16 | 11K, 4R, 1T₀ | 3D, 9E, 1T₉ |
velocities of the system, based on the specific desired temperature (298 K or 310 K). We then applied a harmonic potential with a constant force of 1000 kJ mol\(^{-1}\) nm\(^{-2}\) to all heavy atoms (non-hydrogen) of the protein, restraining their movement for 2 ns, letting the solvent unrestrained to build into a proper solvation layer. Until this step, the thermalization was assisted by the weak coupling scheme, by the Berendsen thermostat [41]. Afterwards, we remove the restraints of the protein, with the bond constraints being controlled by the LINCS [42] algorithm. The cut-off of non bonded interactions was set to 1.0 nm, including periodic boundary conditions. The pressure was maintained constant at 1 atm and the temperature at 298 K or 310 K by the Parrinello–Rahman barostat [43] and by the V-rescale thermostat [44], respectively (NPT ensemble). The time step for algorithm integration was set at 2 fs, and trajectories were saved each 20 ps. Further analyses were performed using GROMACS, VMD [45], PyMol [46], Origin®8 [47] and scripts of our own.

3. Results and discussion

The topology of the DIII is an immunoglobulin C-like module [48]. It consists in five preserved \(\beta\)-strands among all serotypes and has no helices. There is one preserved disulfide bond (Cys302-Cys333) that links the N-terminus of DIII with the first loop after the \(\beta\)-hairpin formed between the \(\beta_1\) and \(\beta_2\) strands (Fig. 1). This bond is remarkable because it links the immediate subsequent portion after the Domain I/III hinge (a region of high mobility and subject to considerable mechanical stress) to a more internal point of DIII. Both this fact and the high conservation of these cysteines throughout the four serotypes indicate that this disulfide bond can be of considerable structural relevance.

Even though the primary structure of the Domain III from all four serotypes differs notoriously (Fig. 1(A)), their tertiary structure is substantially preserved (Fig. 1(B)). Considering its frank exposure to the solvent, such structure conservation is important because DIII may act as the fingerprint of the virus, carrying its specific identification for each serotype as epitope: DIII discriminates the binding sites which form the active complex antibody (Fab 1A1D-2)-protein E (residues from 304 up to 313 and four others, for DENV-2) [49]. Some of the low identity regions present rather significant mutations, for they often alter physicochemical attributes, most of them, notably, in the loops. We built a relationship tree of the four serotypes, with groups DENV-1 and DENV-3 as closely related serotypes, while DENV-2 may be assigned as their sister group (Fig. 1(A)). DENV-4

Fig. 1. Structural alignment from DIII for the four known dengue serotypes. A. DIII alignment of the residues for the four dengue virus serotypes. The coloring follows the Blosum62 table, indicating high to low conservation respectively from darker tones to white. A representation of the correspondent preserved secondary structure is shown (\(\beta_1\) to \(\beta_5\)). On the lower right, the neighbor joining tree reflecting the scores among the serotypes, which are 68 (DENV-1–3), 63 (1–2), 59 (2–3), 58 (2–4), 52 (1–4) and 49 (3–4). Alignment source: ClustalW2 — Multiple Sequences Alignment [50]. B. Final structurally stable configurations of DIII for the four serotypes of DENV. Each was obtained from backbone coordinates, after 200 ns of molecular dynamics, at pH = 7 and 298 K. The common \(\beta\)-strands are sequentially numbered from the N to C terminus. Eventual extra sheets, present in one particular serotype, are identified by a prime signal ('). Note that the \(\beta\)-sheets hold practically the same relative spatial orientation. The four structures were previously aligned, and the fitting is presented on the right. This illustrates the fact that, though the loops float significantly, the first five strands maintain their relative positions.
assumes a more basal position, having a lesser identity with the other three. These relations are in accordance with a previous phylogenetic analysis [51]. These identity scores (Blosum62) reflect and explain some structural behavioral implications, as we present in the following sections.

3.1. General aspects of structural flexibility

The RMSD evolution for each of all four dengue serotypes at pH = 3, 5 and 7, and at two temperatures (298 and 310 K) is shown in Fig. 2. The set of curves illustrates, for each serotype, how the isolated DIII responds structurally to variations on the intensive conditions of the system. The behavior is subtle but specific to each serotype, ranging from a very small structural reaction up to more intense and opposing conformational changes.

Among all cases, the DIII of DENV-1 is the one that shows the smallest conformational changes relative to its original crystallographic structure, as shown by the corresponding RMSD in Fig. 2; it is practically unaffected by variations on the tested intensive conditions. On the other hand, the DIII of DENV-2 suffers larger deformations in respect to the previous case for all tested pH, mainly at the environmental temperature. The other two cases react to very specific conditions: at pH = 7, the DIII of DENV-3 suffers an expressive and maintained deformation, either at 298 K and at 310 K. But, for DENV-4 such deformation occurs at opposite conditions, that is at the smallest pH = 3, and only at the higher temperature of 310 K. The inset of Fig. 2 shows the overall temporal average <RMSD>, which comparatively illustrates the conformational response of DIII to changes in the intensive conditions of the system, for each serotype and temperature. Such average is given by

\[
\text{(RMSD)} = (3 \times N)^{-1} \sum_{(\text{pH})} \sum_{(i)} \text{RMSD}_i \text{(pH)},
\]

that is, <RMSD> is an average taken over all three considered pHs along the entire trajectory, which is composed by a set RMSD, of \( N \) instantaneous values of RMSD. The inset of Fig. 2 indicates that the larger conformational changes occur for DENV-2; for this serotype the DIII departs significantly from its original configuration for all tested intensive conditions. However, the structural difference among the DIII in different pH values is not so substantial and, in this sense, its behavior is somewhat similar to that of DENV-1.

3.2. Local structure flexibility

The series of 200 ns MD simulations enable us to assess the changes on the conformational behavior of DIII in solution as a result of distinct intensive conditions, and to compare those changes among the four serotypes: in general, the isolated DIII can be characterized by a rigid set of \( \beta \)-sheets disposed as an immunoglobulin-like core, flexible loops and, at this particular case, loose termini.

Further details of the DIII conformational behavior are obtained by calculating how the protein’s local flexibility responds to the aforementioned distinct intensive thermodynamic conditions. This physical quantity can be captured by the root mean square fluctuation (RMSF) for each residue: (RMSF), is a temporal average of the displacement of residue “i” (carbon alpha) with respect to its spatial average locus \( \langle \mathbf{r}_i \rangle \), Eq. (2). In essence, the RMSF quantifies the regions of more/less flexibility along the chain (Fig. 3). At a temperature of 298 K, as a general behavior for all serotypes, we note that distinct pH levels affect different regions of DIII in an uncorrelated form: in particular regions, lower (or higher) pH causes the chain to be looser, but in other regions an opposite effect may occur, that is, the same pH level can increase its stiffness. However, at 310 K, the pH level can uniformly perturb the protein as a whole: the entire DIII of DENV-1 becomes less flexible (fluctuates less) at higher pH = 7, while for DENV-2 the behavior is the opposite: DIII becomes more (less) flexible at the higher pH = 7 (lower pH = 3). Interestingly, for DENV-4, DIII becomes

![Fig. 2.](image-url)
extremely rigid at intermediate pH = 5 (T = 310 K) and, as shown in Fig. 3, it is most flexible at low pH = 3.

\[
\text{RMSF}_i = \sqrt{\frac{3B \text{- factor}}{10(8m^3)}}
\]  

(2)

When possible, simulation data should face available experimental values, which is standard for RMSF plots [52]. The preparation of the data followed two protocols. For DENV-1 and DENV-3, which are crystallographic structures, each atom B-factor is provided in the PDB file and, in that case, we employ Eq. (3) to convert the B-factor to RMSF [53]:

\[
\text{RMSF} = \sqrt{\frac{3B \text{- factor}}{10(8m^3)}}
\]  

(3)

On the other hand, DENV-2 and DENV-4, were determined by NMR and do not carry a B-factor information in the PDB file. So we ran an RMSF analysis directly among the models provided in the PDB file. However, one should bear in mind that the models of a NMR determined structure deposited in a PDB file are fit to the experiment, and therefore do not reflect a correct sampling of a Boltzmann distribution in the same way a MD simulation does [54]. The data shows that DENV-1 and DENV-2 fit best between experiment and simulation, and, in spite of a slight quantitative underestimation, qualitatively, the simulations reproduce most of the results. In the case of DENV-4, the experimental data is quantitatively smaller, though the high flexibility of β4–β5 is evident. On the other hand, DENV-3 shows a rather unusual periodic rounded pattern of fluctuation, outlining, at some extent, the high flexibility of the loop β2–β3. It is important to note that even the experimental data of the four serotypes show remarkable distinct profiles on the fluctuation patterns among each other, so the experimental conditions and data collection must be taken into account. Also, the structures which were originally determined as E protein monomer (DENV-1), dimer (DENV-3) or isolated DIII (DENV-2,4), carry different B-factors, for the interaction between monomers and/or domains, may cause significant changes in this variable.

This diversity of conformational behaviors – induced by intensive thermodynamic conditions – results from a complex set of interactions (chemical, electrical and steric) between the side-chains and the backbone, near and far along the protein chain. That means that, a proper mutation of only a few residues can significantly alter the protein’s malleability – locally or abroad – without, however, substantially disrupting

Fig. 3. The root mean square fluctuation (RMSF) for the four dengue serotypes at distinct pH and temperature values. The individual flexibilities of each residue along the chain, as a function of pH (3, 5 and 7) and temperature (298 K and 310 K) variation are plotted. The particularities of pKas of each residue at each given pH generate several ionization configurations and thus the structural variations. The color coding and types of lines are the same as in Fig. 2. Dotted black lines represent the experimental data plotted for comparative purposes. A secondary structure guide is shown in the middle, and represents the average configuration of the correspondent residue in the y-axis.
its basic conformation. Of particular interest here are the changes in the flexibility of loops, which are the most affected regions by changes on the thermodynamic intensive parameters.

The loop involving His317 settled in between sheets $\beta_2$ and $\beta_3$ (loop $\beta_3-\beta_2$) has six residues that are preserved among all four serotypes (Fig. 1); even so, at pH = 3 and temperature of 298 K, the peak flexibility in the DIII of DENV-3 is about 75% larger than that in DENV-4 (Fig. 3). In turn, at 310 K, the flexibility of loop $\beta_3-\beta_2$ at pH = 3 is about the same for all four serotypes, but it can change significantly for other levels of pH. This fact may be illustrated with DENV-1, which when at pH = 3 exhibits flexibility about twice larger than at neutral pH. Similarly, even though the residues of the loop $\beta_3-\beta_4$ are also preserved among all serotypes (Fig. 1(A)), this region reacts diversely under different pHs and temperatures. For instance, while at lower pH = 3 and T = 310 K, the flexibility of the loop $\beta_3-\beta_4$ is maximum for DENV-4 it is minimum for DENV-2. Note that at a temperature of 298 K, this behavior is not reproduced.

On the other hand, a significant fraction of residues of the loop $\beta_4-\beta_5$ and $\beta_2-\beta_3$ are, however, not preserved among the four serotypes and, as a general rule, under the same intensive conditions, their flexibility varies from one serotype to another, in amplitude as much as in tendency. For example, at low pH = 3 and T = 310 K, the small loop $\beta_4-\beta_5$ presents the smallest flexibility in DENV-2, while in DENV-4, the same pH level maximizes it. In the case of loop $\beta_2-\beta_3$, eight residues present the lowest identity among the four serotypes. This larger loop presents the most intense changes on its flexibility under different conditions. Around Cys333, its flexibility is reduced due to a disulfide bond (Cys333-Cys302), and, even so, changes on the pH level can change its local fluctuation profile. The pH level affects more broadly the malleability of $\beta_2-\beta_3$, especially at the temperature of 310 K; in the sequence from Asp341 up to Asn366, the loop flexibility is markedly distinct from one serotype to the other. By comparing pairs of serotypes (for instance, DENV-2 and DENV-4), one can perceive that the amplitude as well as the locus where it is more/less flexible changes profusely.

The fundamental reason for such diverse responses clearly lies on the ionization changes of charged residues under different pH levels: this process can create, remove or alter salt bridges (SB) and hydrogen bonds (HB), as well as alter interactions of chargeable residues with polar atoms. Eventually, such local changes can trigger a chain of complex events, complicating the full description of the process. Particularly, the flexibility of the loops $\beta_2-\beta_3$ and $\beta_4-\beta_5$ are remarkable. This fact was confirmed by analyzing the B-factors of its 20 models in the PDB file (ID: 2SFJ — DENV-2), and may have implications on immunologic recognition, as suggested before[34,14].

In short, the conformational flexibility responds to changes in pH and temperature, both qualitatively and quantitatively, according to each serotype. For DIII as a whole, a few cases illustrate the variability of the response of its structure flexibility in function of the pH level. For DENV-1, the DIII at neutral pH = 7 is the most flexible at T = 298 K, but becomes the most rigid at 310 K; on the other hand, for DENV-2, exactly at pH = 7 and T = 310 K, the DIII behavior is the opposite, that is, it presents the highest flexibility. Particularly for T = 310 K, the local flexibility response to variations in pH is clearly dependent on the serotype; the response can be null, that is: the flexibility is practically the same for pH = 3, 5 or 7 (DENV-3); can be contrasting: at acidic pH = 3, DIII can be extremely flexible for DENV-1 and DENV-4, or extremely rigid (DENV-2); or yet very selective: specifically for DENV-4, the flexibility response of DIII to pH is very well stratified for all residues, being least flexible for pH = 5, most flexible for pH = 3, and presenting a kind of an average flexibility at pH = 7.

3.3. Highlighting the $\beta_2-\beta_3$ and $\beta_4-\beta_5$ loops

The loops $\beta_2-\beta_3$ and $\beta_4-\beta_5$ constitute the sequence or residues with the highest mutation rate of DIII, suggesting specificity to each serotype. They are also regions of the molecule with high flexibilities (Fig. 4), suggesting effective capability of spatial adjustment; they present a high concentration of ionizable residues, including histidines. All these attributes of $\beta_2-\beta_3$ and $\beta_4-\beta_5$ loops and their spatial orientations on the virus surface, facing the five-fold DIII pore at the virus envelope (opposed to the DI/DIII interface), provide evidences of the importance of these particular loops as targets for molecular recognition. The described flexibilities occur at all examined pH values, however particularly for the value of pH = 7, the changes on the chain flexibility on these sites, as a function of DENV type, can be correlated to the selective efficiency of the binding of the virus to the host cell membrane.

Using principal component analysis (PCA)[55], we were able to estimate the main tendencies of molecular motions of DIII (Fig. 5). The results indicate that the $\beta_4-\beta_5$ and $\beta_2-\beta_3$ loops share the same plain of coordinated movement, towards their opposite direction, in a hinge-like basis. Ongoing investigation of this dynamics will reveal more details on the level of interaction between these two structures. Additionally, correlation lines indicate that both DENV-1 and DENV-3 have a long-range structural correlation (more rigid core), in agreement with lower RMSD (Fig. 2). Moreover, we found a high correlation within the loop structures, with interactions ranging from three to as far as ten or more residues. These inner correlations, together with the fluctuation patterns shown in Fig. 3, suggest that some segments of these loops may move essentially preserving their shapes.

The primary structure differences among the serotypes, in spite of the structural maintenance of the core of the DIII by means of the fierce interaction provided by the $\beta$-sheets, significantly shift the

---

**Fig. 4.** Color-coded representation of DIII flexibility (pH = 7, temperature of 310 K). The color coding follows the extent of RMSF: white (≤0.15 nm) represents the rigid core, while light blue (0.16–0.25 nm), purple (0.26–0.35 nm) and dark blue (>0.36 nm) represent regions from light to intense flexibility. Note that most of the higher flexibilities are associated around the loop sections of the protein. The green arrows indicate the DIII termini.
flexible loci of the molecule charges, affecting their superficial electrostatic potential. In spite of the His317, which is ubiquitous, only the DIII of DENV-2 and DENV-3 have an additional His (346 and 347, respectively). They are located in the $\beta_2$-$\beta_3$ loop, and respond to net charge increasing as the environment acidifies. That ionization changes the stereo electrical profile (Fig. 6), and these alterations
could evidently result to important implications in the binding with diverse ligands (such as antibodies) with an induced fit mechanism [56,57], as well as assist in the signaling and response by the host cell.

3.4. Quaternary configuration of DIII: the regions of loops β2–β3 and β4–β5 as an immunogenic site

As seen above, the regions involving the residues 338–346 (β2–β3) and 380–386 (β4–β5) share remarkable stereochemical characteristics. Additionally, they are located facing the five-fold axis of symmetry on the mature virus (Fig. 7), where the DIIIs are closely related, constituting twelve pores throughout the virus surface.

Each one of these pores is slightly more prominent out of the envelope surface [31]; an important topological characteristic that enhances their availability to external interactions, either with antibodies or other types of ligands. Indeed, Zhang and collaborators [14] have already brought attention to the importance of both the 5-fold axis region and three inherent loops (two of them being our referred β2–β3 and β4–β5), all of which share a relatively high temperature factor. Here we set a simulation box with five DIIIs in the pore configuration (details in Section 2.1) in order to discuss in some detail the conformational behavior associated with the electrostatic network in this region, specifically for DENV-2.

Each loop among the domains constituting the pore concentrates high internal HB occupancies, and, in a lesser extent, also inter-loop occupancies (Table 2). These interactions are established only intra-domains, for the distance among the subunits in the loop regions is far too great to promote any HB or SB. RMSD analysis of the loop regions reveals two distinct behaviors: RMSDβ2–β3 presents an average deviation from the reference structure that is almost three times larger than RMSDβ4–β5, being 0.27 ± 0.05 and 0.11 ± 0.02, respectively (Fig. 8(A,B)). Observe that this statement is also true for the standard deviation, which is a direct reflection of the structural fluctuation that arises from the shifting between the two main states of β2–β3, whether it is more narrowed (α1) or widened (α2) (Fig. 8(C–E)). The β2–β3 loop presents a HB pattern that is mostly centered in importance at the aspartic acid residue at the position 341 (Asp341). The main reasons why Asp341 plays a critical role are three. First, it links the basis of the loop by a perennial main-chain/main-chain HB with His346 (average occupancy ϕ = 83.7%); this stabilizes the general fold of this particular secondary structure and its presence narrows the loop, while its absence widens it (Fig. 8(C–E)). Second, its side-chain, specifically the atoms Oδ1 and Oε2, acts as two very efficient acceptors, which fit in a perfect clamping geometry with the donor group NH4+ and NH3+ from the arginine at position 345. This interaction is not present in the crystallographic structure and seems that, once established, it tends to persist. Additionally, Arg345 also interacts intensely with Asp341 through their amide and β-carboxyl group, respectively. Third, Asp341 is 100% preserved among all the sequences of the four dengue serotypes deposited in the Uniprot [58] database (although it is not conserved among other flaviviruses [59]), which makes its importance more evident. As previously stated, Asp341-His346 presents an average occupancy of ϕ = 83.7%, which was calculated considering all five DIIIs in the pore, adding up their individual 200 ns, yielding a 1 μs average. If one considers that the configurational activity of this loop is not significantly affected by the neighborhood, the free energy ΔG related to transition from open to narrow loop, the average rate of occupancy can be estimated by the relation [60,61]

\[
\Delta G = RT \ln(\phi/(1-\phi)),
\]

which yields ΔG ≈ −1.01 kcal/mol, where \( R = 1.99 \) cal/(K mol) is the gas universal constant and \( T = 310 \) K (physiological absolute temperature). At the base of β2–β3, Ile380(NH)-Leu387(O) and 387Leu(NH)-Ile380(O) respond for the highest occupancies (ϕ = 99.26% and ϕ = 96.84%, respectively); hydrogen bonds between Glu383Gly(N)-Val382(O), Gln386(Nε2)-Gly381(O) and Gln386(Nε2)-Glu383(O) are also frequent. Furthermore, interactions between

<table>
<thead>
<tr>
<th>Loop</th>
<th>Pairs [donor-acceptor]</th>
<th>&lt;Occupancy&gt; (%)</th>
<th>ΔG (kcal/mol)</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>β2–β3</td>
<td>Asp341m-His346m</td>
<td>83.7</td>
<td>−1.01</td>
<td>Base connection</td>
</tr>
<tr>
<td></td>
<td>Gly383m-Asp341m</td>
<td>57.4</td>
<td>−0.18</td>
<td>Middle portion connection</td>
</tr>
<tr>
<td></td>
<td>Lys344m-Asp341m</td>
<td>60.8</td>
<td>−0.27</td>
<td>Middle portion connection</td>
</tr>
<tr>
<td></td>
<td>Arg345m-Asp341m</td>
<td>89.4</td>
<td>−1.31</td>
<td>Middle portion connection</td>
</tr>
<tr>
<td></td>
<td>Arg345v-Asp341v</td>
<td>80.4</td>
<td>−0.87</td>
<td>Stabilizes closed state</td>
</tr>
<tr>
<td></td>
<td>Val347m-His346m</td>
<td>37.1</td>
<td>0.32</td>
<td>Adjacent external bond</td>
</tr>
<tr>
<td></td>
<td>Leu348m-His346m</td>
<td>26.1</td>
<td>0.64</td>
<td>Adjacent external bond</td>
</tr>
<tr>
<td>β3–β4</td>
<td>Ile380m-Leu387m</td>
<td>96.8</td>
<td>2.10</td>
<td>Base connection</td>
</tr>
<tr>
<td></td>
<td>Gly385m-Val382m</td>
<td>56.5</td>
<td>−0.16</td>
<td>Middle portion connection</td>
</tr>
<tr>
<td></td>
<td>Gln386m-Gly381m</td>
<td>68.2</td>
<td>−0.47</td>
<td>Middle portion connection</td>
</tr>
<tr>
<td></td>
<td>Gln386m-Glu383m</td>
<td>29.0</td>
<td>0.55</td>
<td>Middle portion connection</td>
</tr>
<tr>
<td>Inter loops</td>
<td>Leu383m-Ile380m</td>
<td>99.3</td>
<td>−3.05</td>
<td>Base connection</td>
</tr>
<tr>
<td></td>
<td>Lys344v-Gln386v</td>
<td>12.2</td>
<td>1.21</td>
<td>β2–β3/β4–β5 connection</td>
</tr>
<tr>
<td></td>
<td>Lys344v-Glu383v</td>
<td>21.4</td>
<td>0.80</td>
<td>β2–β3/β5–β5 connection</td>
</tr>
</tbody>
</table>

Fig. 7. The DIII in the five-fold pore configuration, as found on the native DENV surface. In A, the bottom view of the pore, view from inside the virus. In B, the upper view of the pore, from outside the virus. The black line delimits two DIIIs that were removed in order to produce C, which shows the internal disposition of the loops (note the different mean size of the internal and external diameter of the pore). The β2–β3 loop is colored in red, while the β4–β5 loop is blue. In D we present the icosahedral symmetry of the virus, indicating the location of the five-fold axis pore on the virus surface. The E proteins are placed according to their actual asymmetric unit. The green circumference highlights the pore.
Fig. 8. Details of the electrostatic interactions of DENV-2 inside the five-fold surface pore. A When destabilized around 25 ns, the hydrogen bonds and salt bridges (red bars indicate occupancies) involving Lys344 (as show in F) promote a β2-β3, reconfiguration increasing the distance between both loops (Dβ23), and slightly increasing the pore radius (Dβ45), as demonstrated by d1 and d2 in B, respectively. This reconfiguration highlights the transition from the αC (C–D) to the αs shape (E), with subtle changes in the HB network. Residues that interact through their backbone and/or their lateral chain have the carbons colored in green. Hydrogen bonds and salt bridges are both shown as yellow dotted lines.

β2–β3 and β4–β5 within a subunit are mildly frequent. They are responsible for linking both loops, bringing them closer when salt bridges between Lys344(Nζ2)–Glu383(Oε2) and Lys344(Nζ2)–Gln386(Oε2) are instituted (Fig. 8(F)). This characterizes the aforementioned antagonist movement (hinge-like), as illustrated in Fig. 5, top.

This network dictates the proximity among the loops as well as their basic shape, having also an evident effect on the pore diameter (Fig. 8(A–B)). It also may assist the pore region as a buffer of charges that could become available by reconfiguration of the surface electrostatic potential in the event of an environment change, one that can be driven by newly introduced local triggers, triggered, for instance, by the proximity of another molecule (ligand).

4. Conclusion

The capability of sensing environmental changes has a direct effect on DIII structural fluctuations; such intensity is serotype-dependent and tend to be more pronounced in the β2–β3 and β4–β5 loop regions (Fig. 3). The rigid DIII core is composed mainly of high similarity regions, providing structural stability of the β-sheet that works as a steady foothold to the β2–β3 and β4–β5 loops, which in turn present fingerprint-specificity, and enhanced capability to explore the configurational space. As protein internal flexibility may be required for optimal molecular recognition [62], then the structural behavior of DIII seems to fulfill the conformational attributes to the process of infection.

The low identity regions imply distinct physicochemical properties throughout the DIII among the four serotypes, consequently inducing particular electrical potential patterns on the surface of the virus (Fig. 6), and therefore a specific profile of response to external variation of thermodynamic intensive parameters. The consequences of such subtle local media-induced specificities of each serotype could account, a priori, for one of the responsible facts for the human immune response relative inef-}

References
