

Docking and QM/MM Studies of NS2B-NS3pro Inhibitors: a Molecular Target against the Dengue Virus

Isabella P. Godói,^{*a} William Gustavo Lima,^{a,b} Moacyr Comar Junior,^a Ricardo José Alves,^c Jaqueline Maria S. Ferreira,^b De-Xin Kong^d and Alex G. Taranto^a

^aLaboratório de Química Farmacêutica Medicinal and ^bLaboratório de Microbiologia, Universidade Federal de São João Del-Rei, 35501-296 Divinópolis-MG, Brazil

^cLaboratório de Química Farmacêutica, Departamento de Produtos Farmacêuticos, Faculdade de Farmácia, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte-MG, Brazil

^dCenter for Bioinformatics, College of Life Science and Technology, Huazhong Agricultural University, 430070 Wuhan, P. R. China

Dengue virus (DENV) has been characterized as having great clinical importance in the world, as there is no specific treatment against this virus. The NS2B-NS3pro complex is essential for the replication and maturation of DENV and is a potential pharmacological target. The present study aims to evaluate and understand the interactions and affinities (via molecular docking/AutoDock Vina) of 16 peptidomimetic derivatives applied to a NS2B-NS3pro DENV-2 complex constructed by homology modeling (via SWISS-MODEL). Two compounds were selected as potential inhibitors of this protein complex. In addition, these compounds possess important interactions involving Ser135, Gly169 and Tyr161, which have been described previously to be fundamental to the recognition of inhibitors directed to this receptor. Thus, the involvement of these residues is significant pharmacologically because they may contribute to the inhibitory action of this molecular target against DENV.

Keywords: dengue, inhibitors, NS2B-NS3pro, molecular modeling

Introduction

Dengue fever affects approximately 390 million people annually, with higher prevalence in tropical and subtropical regions.¹ Dengue virus (DENV) belongs to the genus *Flavivirus* and *Flaviviridae* families and is transmitted by the bite from female mosquitoes of the genus *Aedes* and species *A. albopictus* and *A. aegypti*.² Currently, there is no specific therapy available against DENV; however, palliative treatments to circumvent symptoms caused by the disease are available. DENV has four antigenically distinct serotypes (i.e., DENV1-4), which makes it difficult to develop an effective vaccine.²⁻⁴

The viral RNA encodes 10,173 nucleotides that are translated into a polyprotein. This protein is then cleaved by viral and cellular proteases, culminating in the release of three structural proteins (i.e., capsid (C), precursor membrane protein (prM) and envelope (E)) and seven nonstructural

proteins (i.e., NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5).^{2,5} The serine protease called NS2B-NS3pro DENV has considerable genetic similarity with the analogous West Nile virus (WNV) protease.⁶ This protease complex is directly related to replication and maturation of the virus, and it thus represents a promising molecular template for the treatment of these two pathogens.⁵⁻⁷

The N-terminal region of NS2B-NS3pro has 187 amino acids and the catalytic triad His51-Asp75-Ser135 constitutes the active domain of the serine protease, which is essential for cleavage of the viral polyprotein.^{6,8} Its catalytic mechanism involves the nucleophilic attack of the oxygen from the serine hydroxyl group and the electrophilic carbon of the peptide bond to be cleaved.^{6,9-11} Erbel *et al.*⁶ identified important residues for inhibitor recognition in NS3pro such as Gly151, Gly153 and Tyr161. Regarding the organization of the flaviviral active site, Aleshin *et al.*¹² demonstrated that a precise spatial arrangement is required to enhance the nucleophilicity of the serine hydroxyl group present in the His-Asp-Ser catalytic triad. Furthermore, the “oxyanion

*e-mail: labqf_ufsj@ufsj.edu.br

hole" is stabilized by Ser hydroxyl H-bonds to the His imidazole, which in turn H-bonds to the Asp carboxyl group.¹² Consequently, a tetrahedral intermediate is formed and stabilized by hydrogen bonds involving Gly151 and Ser135, which comprise the oxyanion pocket.¹⁰ Thus, an inhibitor that is able to interact with these amino acid residues represents a potential candidate for the inhibition of the serine protease. The positions of the inhibitor P1, P2, P3 and P4 are related to the pockets S1, S2, S3 and S4 in this receptor, respectively.^{6,11-13}

Pharmacologic therapy is absent for Dengue fever. Several research groups have begun medicinal chemistry projects using rational approaches. Thus, drug development efforts have increased considerably.^{11,13-18} Theoretical approaches in this regard include several methodologies, such as homology modeling,¹⁹ virtual screening,^{20,21} molecular docking,²²⁻²⁵ molecular mechanics (MM) and quantum mechanics/molecular mechanics (QM/MM).²⁶⁻³³ Homology modeling, or comparative modeling, aims to build three-dimensional models of a protein (target). The target sequence is compared with a similar protein (template) whose structure has been determined experimentally.¹⁹ Then, a structural alignment tool is used to construct the model.²⁷ After the model is developed, refined and evaluated, methodologies such as molecular docking and virtual screening can be performed to determine new lead compounds.^{21,23,28,30}

Generally, the 3D structure of a molecular template directs the molecular docking simulations.²¹ Docking methodologies consist of a computational resource that attempts to predict the non-covalent interactions between the macromolecules (receptors) and small molecules (ligands) at the active site. Further, they aim to evaluate the bound conformations and binding affinities. As a result, docking is able to rank the ligands through electrostatic and van der Waals forces.^{20,22} Currently, this methodology is considered to be a high performance tool to search for an appropriate ligand in a set of thousands of compounds (virtual screening, VS)³⁴ or to find a molecular template for a specific ligand (inverse virtual screening).³⁵ Such docking simulations can be performed using rigid or flexible approaches. In the former case, all rotatable bonds of the molecular template are kept rigid, whereas the rotatable bonds of the ligand are free to adjust to the binding site. In contrast, flexible docking permits the side chains of selected amino acids and the ligand to be fully rotatable. In general, docking simulations are performed using the rigid approach. Therefore, as expected for a screening tool, false positives can frequently occur during the simulations.²² Thus, more accurate methodologies, such as molecular dynamics simulations (MD) or combined

quantum mechanics and molecular mechanics (QM/MM), must be used to refine and evaluate the docking results to obtain more reliable results.^{30,36}

Another challenge for molecular docking methodologies concerns predicting the role of water molecules at the protein-ligand interface, which are neglected in most protocols.³⁷ The water molecules can increase the stability of the system and act as a competitive agent in the formation of the ligand-receptor complex. Hence, molecular docking simulations with water molecules in the active site may also lead to more reliable results.^{37,38}

This paper aims to construct a 3D model of NS2B-NS3pro DENV-2 for the analysis of potential inhibitors belonging to the class of peptides that possess a Bz-Nle-Lys-Arg-Arg-H prototype inhibitor and other compounds that may interact with the receptor constructed for DENV-2.

Methodology

Comparative modeling

Initially, the structure of DENV-2 complexed with an inhibitor was constructed by the homology modeling method.¹⁹ A 3D model of the NS2B-NS3pro DENV-2 complex was constructed using the Deep View software³⁸ using two three-dimensional structures of NS2B/NS3 serine proteases available in the Protein Data Bank (PDB): DENV-2 (2FOM) and WNV (2FP7).⁶ We note that the latter protease is complexed with the inhibitor tetrapeptide Bz-Nle-Lys-Arg-Arg-H.¹⁵

The alignments of these sequences were performed using the CLUSTALW software program.³⁹ The atomic coordinates of the peptide inhibitor Bz-Nle-Lys-Arg-Arg-H and 13 water molecules were transferred from 2FP7 to construct a model of DENV-2.^{6,19} The water molecules were selected from the molecular template using a radius of 8 Å from the residues of the catalytic triad. Then, the model was optimized by MM using the ff03ua force field in a generalized Born model.⁴⁰⁻⁴³ The model was optimized by MM in 2,000 and 13,000 cycles using the steepest descent and conjugate gradient algorithms, respectively. A cutoff value of 14 Å was used. All calculations were performed using the Sander algorithm implemented in AMBER 11.⁴⁴⁻⁴⁶ Finally, the quality of the model was evaluated via the PROCHECK software program.⁴⁷

The change in the numbering shown for the catalytic triad of the constructed NS2B-NS3pro DENV-2 model compared with that described in the literature (i.e., His57, Asp102, Ser195 to chymotrypsin) results from the different compositions and sequences of these serine proteases (His51-Asp75-Ser135).^{6,9}

Molecular docking

The re-dock calculation was performed using the three-dimensional model NS2B-NS3pro DENV-2 with some associated water molecules.³⁷ Calculations of the root-mean-square deviation (RMSD) and the 2D pharmacophoric diagram of the receptor-ligand interactions were performed by the Discovery Studio Visualizer 3.1 software program.⁴⁸ In addition, the presence of selected amino acid residues in the active site of this molecular template and related residues in WNV⁶ were identified through their interaction between D1 tetrapeptide and NS2B-NS3pro DENV-2.

The structures of the investigated peptides (Figure 1) were obtained from a previous report¹¹ and generated using the GaussView 5.0 software program.⁴⁹ Compounds with basic functional groups were kept in their protonated forms.⁵⁰ These peptides were refined by the AM1 semi-empirical method as implemented in the Gaussian 9 software program. Biological systems can be satisfactorily evaluated using the semi-empirical AM1 method, because it uses Gaussian functions that circumvent problems associated with its design, such as the evaluation of hydrogen bonding. Furthermore, it has experimental parameters for the elements H, C, N, O, F, Al, Si, P, S, Cl, Zn, Ge, Br, I, and Hg.⁴⁹⁻⁵¹

A rigid molecular docking methodology was performed to determine the ligand-receptor affinity and the pharmacophore conformation. The region for the docking process was established using a grid box centered on the catalytic triad His51-Asp75-Ser135. The region was chosen such that it could enclose the entire binding site and the evaluated compounds (i.e., coordinates $x = 20.077 \text{ \AA}$, $y = 43.838 \text{ \AA}$, $z = -0.873 \text{ \AA}$ with the edge set to 1 \AA and $30 \times 24 \times 30 \text{ \AA}$ by the AutoDock Tools). Subsequently, molecular docking was performed using AutoDock Vina 1.5.4.^{22,24} To refine the three-dimensional model, all complexes obtained by molecular docking were minimized using the same parameters described previously.⁴⁰⁻⁴² Finally, the calculations of the relative energies from the molecular docking and QM/MM calculations were performed by subtracting the value of D1 from the complex energy for each peptide (Relative energy = $E_{\text{peptide}} - E_{\text{D1}}$).⁵² Thus, negative and positive values indicate smaller and higher affinities, respectively, for NS2B-NS3pro DENV-2 compared with D1.

QM/MM minimization

QM/MM methods are an important tool to evaluate compounds of pharmaceutical interest because they

combine the accuracy of quantum chemistry with the speed of and molecular mechanic, thereby minimizing the overall computational costs for each independent analysis.^{31,53} However, this hybrid method requires a force field that is specific for the QM portion and it must be sufficiently flexible so as not to polarize the MM regions, which might compromise the integrity of the simulations.⁵³

All complexes obtained by molecular docking were refined using the Sander algorithm as implemented in the AMBER software package and employed the steepest decent (SD) and conjugate gradient (CG) algorithms.⁵⁰ The implicit solvent generalized Born (GB) model, with the IGB flag equal to 1, was selected to perform the QM/MM calculation. This method uses standard equations for the GB model to solve the Poisson equation using the default radii established by LEaP.^{54,55} The molecular mechanics calculations used the ff03ua force field^{40,41} and the semi-empirical method AM1⁵⁰ for the quantum mechanics calculations.^{28,30}

The QM/MM minimization was performed for 5000 cycles using the CG algorithm. The ligand and the rest of the molecule were defined in QM and MM regions, respectively. The cutoff values used for the interaction of the atoms not contained in the MM and QM/MM regions were set to 14 and 20 \AA , respectively. In addition, the conserved water molecules of the binding site were kept present, forming a hybrid solvent model. On the basis of the QM/MM simulations, the relative energy of each compound was calculated using $E = E_{\text{complex}} - E_{\text{protein}} - E_{\text{ligand}}$.⁵²

Results and Discussion

Comparative modeling

The CLUSTALW program³⁹ indicates that the sequence alignment between 2FOM and 2FP7⁶ for the target and template proteins share 45% sequence identity (Figure 2). We note that the amino acid residues present in the catalytic triad are conserved in both structures. Because the threshold value for the construction of a model is 30%,¹⁹ the 45% sequence identity between 2FOM and 2FP7 indicates that the generated model is amenable for molecular docking studies. Additionally, the binding site remains conserved, as highlighted in Figure 2. The His51-Asp75-Ser135 catalytic triad together with Gly151 are important residues present in the oxyanion pocket.^{6,10}

The RMSD value is 0.80 \AA for the overlap of the α -carbons of the NS2B-NS3pro complex obtained from 2FOM (DENV-2 without ligand), 2FP7,⁶ and the three-dimensional model constructed for DENV-2 (Figure 3). This value indicates a good superposition between the 3D

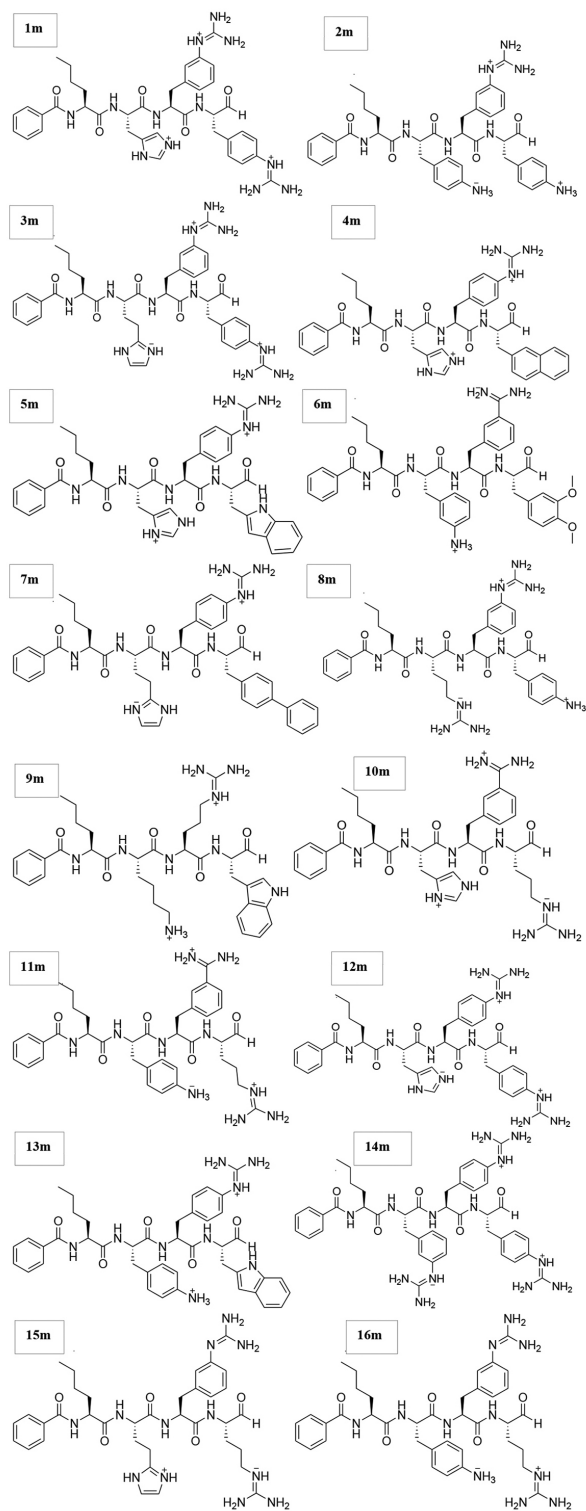


Figure 1. Selected inhibitors (1m-16m) for N2SB-NS3-pro DENV-2.¹¹

model obtained for docking studies, which also suggests a conformational similarity between it and the model.

A Ramachandran plot was used to evaluate the stereochemistry of the DENV-2 3D model, as shown in Figure 4. Figure 4A shows that 82 and 14% of the

residues are the most favored and allowed conformations, respectively. Only 2.8 and 1.4% of the 187 amino acids in the present conformation are tolerated and unfavorable, respectively.⁴⁷ The residues in the unfavorable region (i.e., Gln35 and Glu91) are not present in the binding and catalytic site of this protease, which may arise from regions not being conserved between the available primary sequences. In addition, each residue contributes differently to the tridimensional structure in a manner that is directly related to factors such as stereochemistry, which can affect the torsions shown in the 3D model. Hence, these results do not affect the consequent interactions of this molecular template. The quality of the Ramachandran plot generated for the DENV-2 model was validated by evaluating the percentage of the protein residues that are in more favorable regions. We note that the quality was within the range expected by the internal quality control PROCHECK program (Figure 4C). To compare the model and crystallographic structures, a Ramachandran plot was generated for the 2FP7 structure as well. The Ramachandran plot of 2FP7 (Figure 4B) shows 92 and 8% of the residues are in the most favored and permitted conformations, respectively. Most of the residues present in the model were constructed with the lowest energy conformation, which is likely to be more realistic in a biological environment. Finally, we note that residues located in the unfavorable regions are predominantly arranged in loop portions, which have little influence on the catalytic properties of the evaluated target.

Using the Verify 3D software, an analysis of the environmental preference for each residue showed that 89.59% of the residues had an average 3D-1D score greater than or equal to 0.2 (Figure 5). Furthermore, the model showed the compatibility of the atomic model (3D) with its own amino acid sequence when the results were compared with established structures.⁵⁶

Molecular docking

Initially, the molecular docking was evaluated with re-dock between D1 and the three-dimensional DENV-2 model in the presence of water molecules.³⁷ From the overlay of the re-docked structure and the X-ray crystal structure, the RMSD value was evaluated to be 0.73 Å. This result validates this molecular docking method as it represents a low conformational deviation of the prototype inhibitor structure deposited in the PDB.⁶ Figure 6A shows the re-dock study with the NS2B-NS3pro DENV-2 model and the inhibitor Bz-Nle-Lys-Arg-Arg-H, D1 (Figure 6B).⁶ As shown, the small conformational deviation occurred mainly in the amino acid side chains


```

2FP7_B|PDBID|CHAIN|SEQUENCE  -----GDTTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLW 35
2FOM_B|PDBID|CHAIN|SEQUENCE  AGVLWDVPSPPVVGKAELEDGAYRIKQKGLGYSQIGAGVYKEGTFHTMW 50
                                .:  *.*.*  :.*.*  *  *.*.*  *.*.*.*
                                .:  *.*.*.*  :.*.*  *.*.*  :.*.*  :.*.*.*

2FP7_B|PDBID|CHAIN|SEQUENCE  [H]TTKGAALMSGEGRLDPYWGSKED[RL]CYGGPWKLQHKWNGHDEVQMIIV 85
2FOM_B|PDBID|CHAIN|SEQUENCE  [H]VTRGAVLMHKGRRIE[PSWADVKK]LISYGGGWKLEGEWKEGEEVQVLAL 100
                                *.*.*.*  *.*.*.*  *.*.*.*  :.*.*  *.*.*  :.*.*  :.*.*.*

2FP7_B|PDBID|CHAIN|SEQUENCE  EPGKNVKNVQTKPGVFKTPEGEIGAVTLDYPTGT[SSPIVDKNGDVI]GLY 135
2FOM_B|PDBID|CHAIN|SEQUENCE  EPGKNPRAVQTKPGLFKTNTGTIGAVSLDFSPGT[SSPIVDKKGKVV]GLY 150
                                *.*.*.*  :  *.*.*.*.*  *  *.*.*.*  :  *.*.*.*.*  :.*.*.*.*

2FP7_B|PDBID|CHAIN|SEQUENCE  GNGVIMPNGSYISAIVQGERMEEPAPAGFEPEMLRKK 172
2FOM_B|PDBID|CHAIN|SEQUENCE  GNGVTRSGAYVSAIANTEKSIEDNPE-IEDDIFRK- 185
                                *.*.*.*  :.*.*.*.*  :  *  *  *  *  :  :.*.*.*

```

Figure 2. Sequence alignment between 2FP7 and 2FOM using the CLUSTALW program.³⁹ The colors indicate the following types of residues: hydrophobic = blue; hydrophobic tendency = light blue; basic = red; acidic = purple; hydrophilic = green; unconserved = white. The highlighted residues are His (H), Asp (D) and Ser (S) present in the catalytic triad and Gly (G) in the oxyanion pocket.⁶

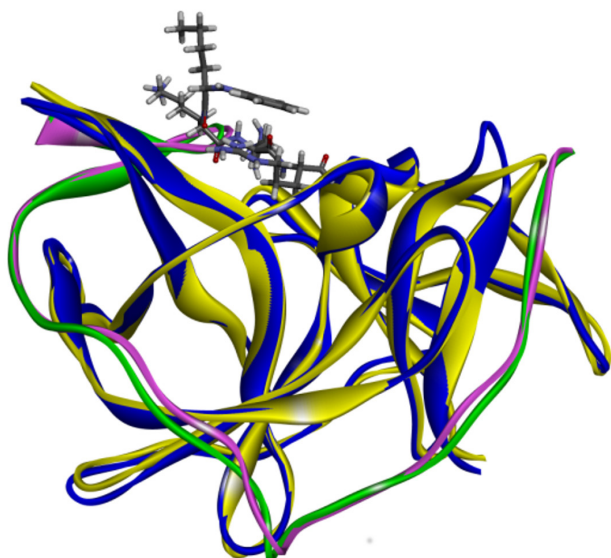


Figure 3. The overlap between 2FP7 (WNV) and DENV-2 model structures complexed with the Bz-Nle-Lys-Arg-Arg-H inhibitor (ribbon representation). 2FP7: NS2B green, NS3 yellow; DENV-2 model: NS2B pink, NS3 blue.

(e.g., phenyl and guanidine moieties). After the re-dock analysis, we identified certain amino acid residues through the interaction between D1 and NS2B-NS3pro DENV-2 (S1: His51, Asp75, Ser135, Gly151, Tyr161; S2: Asn152; S4: Val154) that were consistent with the active site of NS2B-NS3pro WNV,⁶ as shown in Figure 7B.

The water molecules located near the active site in the constructed 3D model conserved their positions relative to the crystallographic structure of NS2B-NS3pro (PDB:2FP7).⁶ However, the ligand adopts a different conformation, which appears to be caused by hydrogen bonds and steric effects. Consequently, these effects elicit differences in the binding energy. In some cases, water molecules act as intermediary agents between the ligand and molecular template.³⁷ Alternatively, these molecules can stabilize the ligand-receptor interactions via hydrogen bonds. In addition, the inclusions of conserved water

molecules increase the prediction of docking pose from 39 to 69%. Thus, we chose to follow this protocol to obtain more accurate results.⁵⁷ Clearly, water molecules must be considered in computational studies directed toward drug development. The complex formed by the docking procedure was subjected to a QM/MM^{31,39} optimization to increase the accuracy of the results. Table 1 shows the binding energy (kcal mol⁻¹) of each compound, which ranged from -13.80 to -7.50 kcal mol⁻¹ for the molecular docking studies and from -568.3 to -18.58 kcal mol⁻¹ for the QM/MM calculations. The binding energies of the complexes help to identify inhibitor candidates for NS2B-NS3pro DENV-2, the molecular template and thirteen water molecules near the active site that were removed from the crystallographic structure (NS2B-NS3pro, PDB:2FP7).⁶ Additionally, the charged nature of some species in this work and the employed semi-empirical method contributed to the results for the complexes obtained with the QM/MM calculations.⁵⁸⁻⁶¹ In this context, Tripathi *et al.*⁵¹ investigated the performance of the charge model in the prediction of docking poses against CDK2 proteins with their respective inhibitors using different semi-empirical (i.e., RM1, AM1, PM3, MNDO) and *ab initio* (HF) and DFT methodologies. The authors demonstrated that AM1 was a useful charge model for the design of new drugs and demonstrated that the scoring function based on this method appears to be applicable in drug design. In addition, the difference in the relative binding energies for the same set of compounds can be understood in terms of atom types and charge effects. AutoDock Vina uses its own atom types and Gasteiger charges. However, in the present work the QM/MM calculations used the AM1 methodology for the ligand and the ff03ua force field for the whole protein. In the QM region, AM1 performs the calculations without atom types and uses “am1-bcc” as the charge methodology. The ligands have charges of 3+ and 2+, which, in effect, are diffuse in the QM region. Overall, these features can contribute to the different relative binding energies.

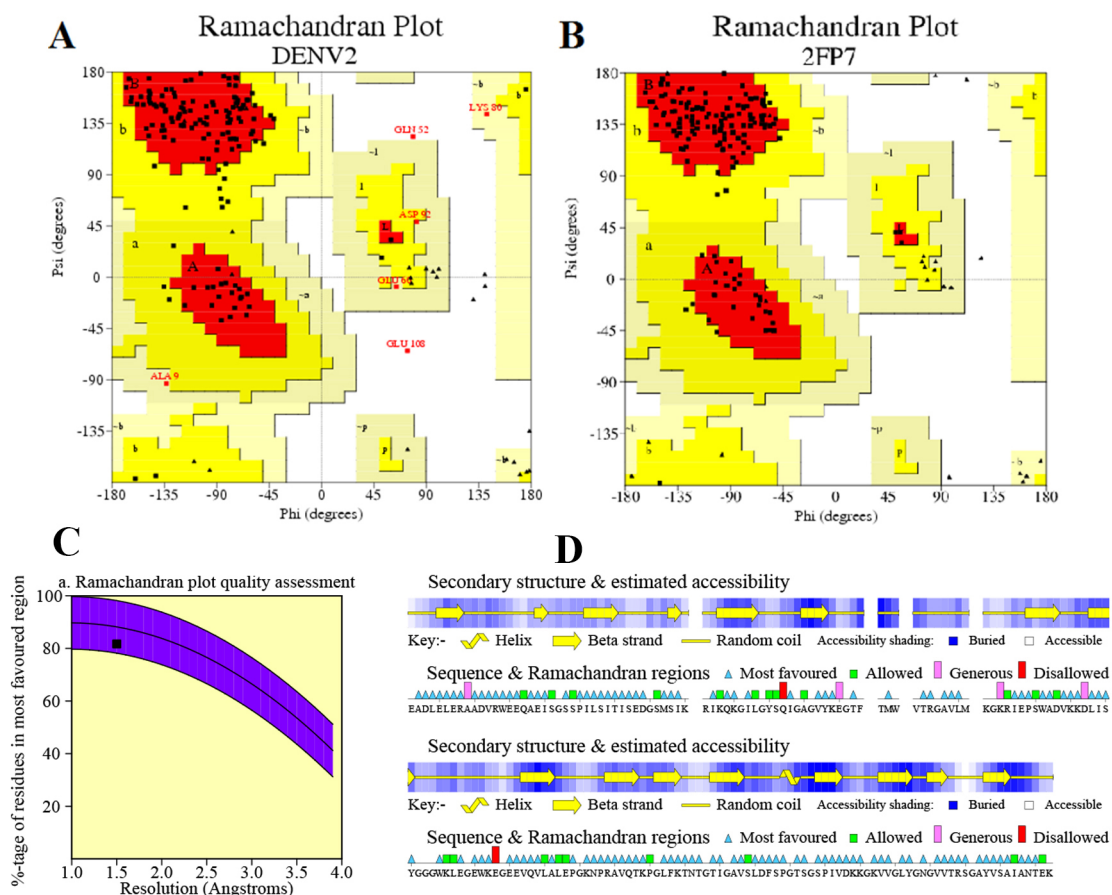


Figure 4. Stereochemical quality of the DENV-2 model and 2FP7. Ramachandran plots of the DENV-2 model (A) and 2FP7 (B) generated by PROCHECK. Red indicates the most favored regions; yellow indicates allowed regions; beige indicates tolerated regions and white indicates unfavorable regions³⁸. Percentage of residues that are most favored in the Ramachandran plot of DENV2 (C). The dark band represents the results of the protein structure; the central line is a mean trend and the width of the band corresponds to the mean standard deviation. Residues of maximum deviation from ideal values (red rectangles). These residues are shown in the primary and secondary sequences and in the region of the Ramachandran plot in which they are located.³⁸

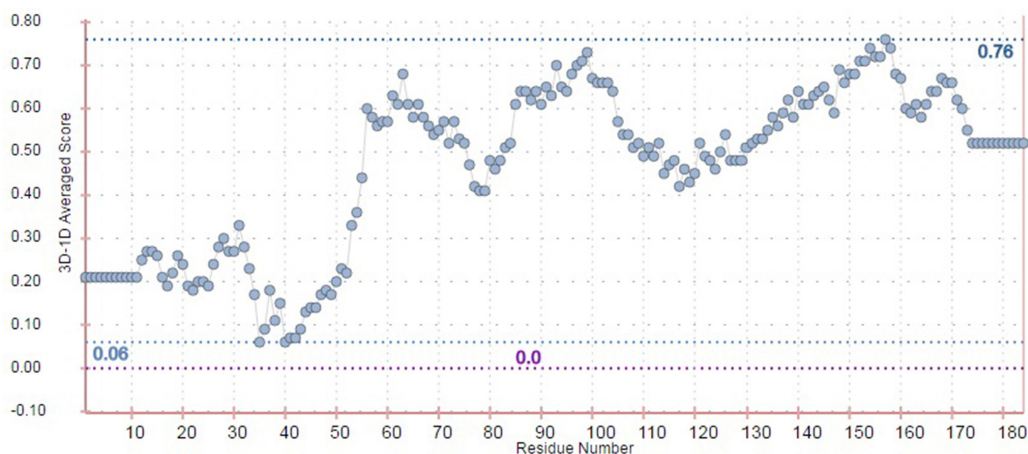


Figure 5. Analysis of NS2B-NS3pro DENV-2 using Verify 3D.⁵⁶

After the combined analysis, D1, **6m** and **12m** were observed to possess high affinities for NS2B-NS3pro DENV-2. The intermolecular interactions were evaluated by a 2D pharmacophoric diagram, with a 2.5 Å cutoff for hydrogen bonds.⁴⁸ The affinity energies of D1, **6m** and **12m**

obtained in the molecular docking procedure are -13.80 , -9.10 and -8.80 kcal mol⁻¹, respectively. QM/MM studies gave affinity energies for these compounds of -568.30 , -180.16 and -138.85 kcal mol⁻¹, respectively. Ligand **6m** gave the best result for the molecular docking and

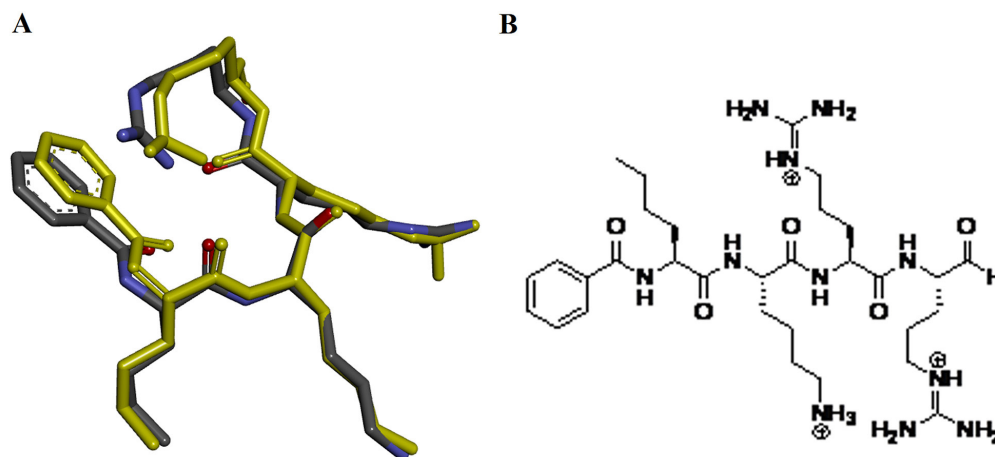


Figure 6. (A) Re-dock stick representation of the Bz-Nle-Lys-Arg-Arg-H inhibitor; (B) the structure of D1.^{6,48} Yellow and gray represent the crystallographic and docked structures, respectively.

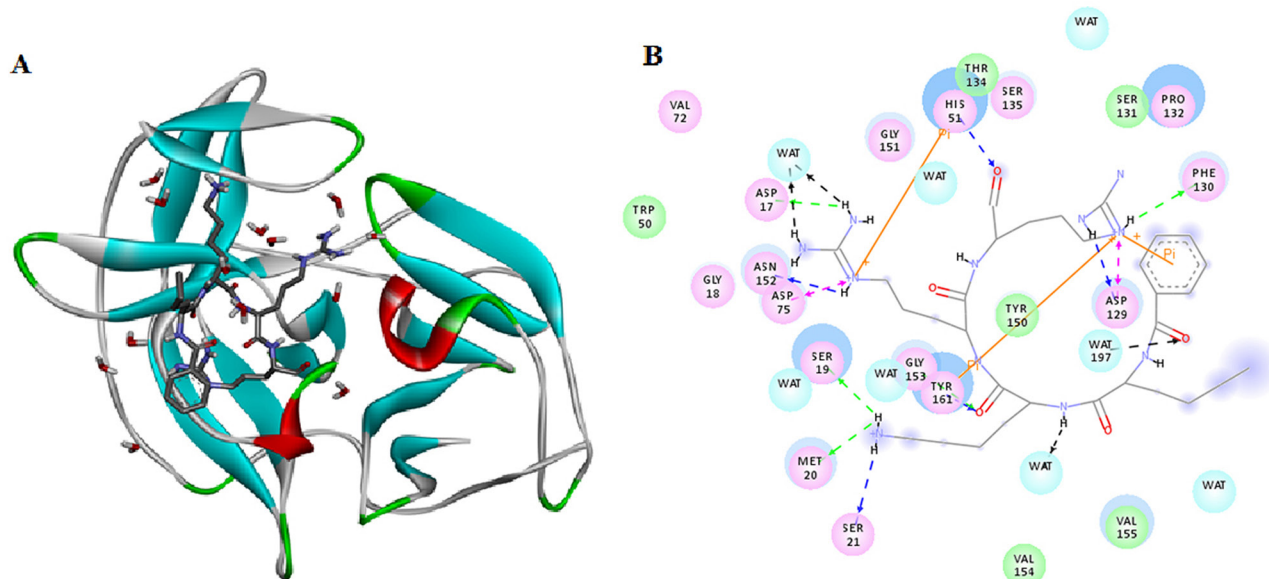


Figure 7. (A) 3D structure of D1 with NS2B-NS3pro and water molecules; (B) interactions of D1 with the active site of this receptor, showing hydrogen bonds (pink circles with traced lines), hydrophobic interactions (green circles), pi-interactions (orange lines), water molecules (blue circles) and close contact residues (pink circles only). The solvent accessible regions are shown as purple circles (in the molecule) and a blue halo (in the amino acid residue). This figure was construed utilizing Discovery Studio Visualizer 3.1.⁴⁸

QM/MM minimization procedures. This ligand also showed a C–O Ser135 distance closer to D1, which increases its interaction with Ser135. In contrast, some compounds showed smaller affinities in both approaches, such as **1m** and **8m**, which possess affinity energies of -7.5 and -7.6 kcal mol⁻¹ (molecular docking) and -75.30 and -52.50 kcal mol⁻¹ (QM/MM). The larger distances of 17.39 and 12.88 Å for the interaction of **8m** and **1m** to Ser135, respectively, in the QM/MM studies can be associated with the absence of hydrogen bonds involving these compounds, which are present in D1. Additionally, **1m** and **8m** possess a low number of binding interactions with the receptor residues, which can help rationalize their smaller binding affinities compared with D1, **6m** and **12m**. There are many

factors that can impact a ligand's binding affinity, such as its structure, stereochemistry, steric features and the number and type of binding (i.e., hydrogen or hydrophobic). For the compounds examined in this work, the number and type of binding can be used to understand the different binding behaviors.

The amino acids that constitute the binding site of the NS2B-NS3pro complex are essential for its interaction with tetrapeptide inhibitors. Some of these residues, such as His51, Asp75 and Ser135, are directly related to enzymatic activity.^{6,10,12} Previously, the residues Gly151 and Try161 have been described to be being important in recognizing inhibitors directed to this receptor. The C–O Ser135 distance, which involves the carbon atom of the aldehyde

group near the hydroxyl group of Ser135, is believed to be an important moiety to enable the catalytic activity of serine proteases.⁶

Following the molecular docking and QM/MM minimization procedures, the interactions involving the inhibitor prototype D1 and the compounds **6m** and **12m** were used as the QM region of the complex, which showed satisfactory results for both methods. Figure 7B shows the interactions observed for D1 within the active site of NS2B-NS3pro DENV-2. D1 possesses hydrogen bond interactions with three residues in this catalytic triad protease (i.e., His51, Asp75, Ser135).⁶ Hydrogen bonds were also observed with Asp17, Gly18, Ser19, Met20, Ser21, Asp129, Gly151, Asn152, and Tyr161. D1 also possesses three cation- π interactions with the Tyr161 and His51 residues. Finally, five hydrophobic interactions were found between the ligand and the residues Trp50, Ser131, Thr134, Tyr150, Val154 and Val155.

Molecular interactions between the ligands and the receptor are related to biological activity. A molecule's structure has a direct influence on its physical and chemical properties, and its geometry can determine its affinity with the receptors in aqueous (physiological) solution (e.g., the effect of positively charged groups present in the inhibitor candidates for serine protease). Additionally, the molecular structure of the receptor determines which moieties in the drug are essential for binding, such as

negatively charged groups (e.g., Asp) present in the active site of NS2B-NS3pro DENV-2.⁶² Figure 8 shows the interactions observed for compounds **6m** and **12m**, which demonstrated higher affinity energies when complexed with NS2B-NS3pro DENV-2. Similar to D1, compound **6m** in the P2 position is involved in two hydrogen bonds with Phe113 (2.1 Å), which are present in the active site of the enzyme. The ligand-receptor association can be enhanced by hydrophobic interactions, as demonstrated by the residues Ser16, Gln18, Ile19, Arg37, His51, Pro115, Gly116, Gly144, Asn152 and Tyr161.

Compound **12m** demonstrated van der Waals interactions with Ser17 and Ser131 and other interactions with important amino acids present in the catalytic cavity, such as π -interaction with His51 and hydrogen bonds with Tyr144 (2.5 Å), which indicates an important relationship to the inhibitory mechanism for this class of enzymes. Additionally, compounds **6m** and **12m** possess a large number of close contact residues (four and eight, respectively), which add further stabilization and are indicative of complementary binding within the activity site. Compounds **1m** and **8m** demonstrated the poorest results, especially for the molecular docking affinity studies and the C–O Ser135 distance. Following the docking study, **1m** and **8m** possessed the smallest number of hydrogen bonds, with one interaction each (Supplementary Information). Compound **1m** showed hydrogen bond

Table 1. Binding energies (kcal mol⁻¹) and C–O Ser152 distances (Å) from the docking and QM/MM procedures

Compound	Molecular docking / (kcal mol ⁻¹)	Relative energy	QM/MM Minimization ^a / (kcal mol ⁻¹)	Relative energy	Distance C–O Ser152 / Å
D1	-13.80	0.0	-568.30	0.0	3.38
1m	-7.50	6.30	-5.30	493.00	12.88
2m	-8.40	5.40	-125.74	442.56	10.31
3m	-7.70	6.10	-48.68	519.62	8.81
4m	-8.00	5.80	-82.62	485.68	10.98
5m	-7.90	5.90	-84.14	484.16	11.57
6m	-9.10	4.70	-180.16	388.14	8.63
7m	-8.10	5.70	-89.12	479.18	9.32
8m	-7.60	6.20	-52.50	515.8	17.39
9m	-7.70	6.10	-57.91	510.39	10.60
10m	-8.10	5.70	-61.07	507.23	8.96
11m	-7.50	6.30	-18.58	549.72	8.88
12m	-8.80	5.00	-138.85	429.45	8.87
13m	-8.10	5.70	-44.17	524.13	10.27
14m	-7.90	5.90	-66.11	502.19	10.02
15m	-8.20	5.60	-63.67	504.63	8.92
16m	-8.00	5.80	-78.00	490.30	8.71

^aBinding energy = $E_{\text{complex}} - E_{\text{protein}} - E_{\text{ligand}}$.⁵²

interactions involving only Tyr161 and van der Waals interactions with Ser34, Pro102, Pro115, Phe130, Ser131, Gly159, Ala160 and Asn162. Following the QM/MM studies, **8m** was involved with hydrogen bond interactions with Gly153 and van der Waals interactions with Ser131, Pro132, Thr134, Ser135, Tyr150, Gly151, Asn152, Val155, Tyr161 and Ser153. The fewest number of hydrogen bond interactions of both compounds and the small quantity of close contact residues (two for **1m** and three for **8m**) contribute to the lower binding energy of these compounds relative to compounds **6m** and **12m**.

The structure-activity relationship (SAR) for the prototype D1 tetrapeptide inhibitor indicates four positions that are responsible for molecular recognition. The positions of the inhibitor denoted by P1, P2, P3 and P4 are directly related to the binding pockets S1, S2, S3 and S4 in NS2B-NS3pro for both DENV and WNV.⁶ S1 is formed by residues Asp129, Tyr 150, Gly151, and Tyr161, which stabilize the positive charge of the arginine side chain in P1. The S2 pocket is negatively charged, which may originate in the backbone carbonyl oxygen atoms of the NS2B residues, such as Asp81. In this way, the positively charged guanidinium group of the P2 arginine is important in forming hydrogen bonds with the negatively charged amino acids of the S2 pocket. The S3 and S4 pockets have been suggested to possess polar (i.e., Leu70, Phe68 and Gln69) and hydrophobic (i.e., Val137) residues, respectively.¹¹

In general, the inhibitor candidates for NS2B-NS3pro DENV are positively charged (3+ or 2+) and possess an electrophilic aldehyde moiety in the P1 position. The most active compounds have this structural profile and can bind residues such as Asp129 (electrostatic), Tyr150 and Tyr 161 (cation- π), which are present in the S1 pocket to form interactions with **6m**, as shown in Figure 8. These results are consistent with those found by Erbel *et al.*,⁶ who revealed key interactions for substrate recognition. These interactions involved the positively charged side chain of arginine (P1 position of inhibitor), which is present in ligand D1, with the Asp112 residue. In addition, the negative electrostatic potential of NS2B residues 82-84 was highlighted. These residues were found to stabilize the guanidinium group of the arginine residue in the P2 position of the inhibitor, which can also form hydrogen bonding interactions with NS2B Gly66 and NS3pro Asn135.⁶

The presence of positive charge in the P1 and P2 portions is not the only important consideration that must be evaluated for inhibitor candidates of this molecular template. Indeed, steric effects, the identity of the *meta* and *para* positions, and the participation of water molecules

must also be considered. Each compound has characteristics that enable a higher or lower binding affinity with the receptor. For example, in both the docking and QM/MM methods, **12m** possessed a lower binding affinity compared with **1m**, suggesting the importance of steric and electronic effects in determining their distinct intermolecular interaction profiles. The number of interactions involved with the receptor can contribute to the greater binding affinity of **6m** and **12m**. These compounds demonstrated the highest number of hydrogen bonds compared with the other investigated compounds, with ten and six interactions, respectively. Furthermore, the Mulliken charges obtained by the semi-empirical AM1 method showed that the most active compound has a partial charge of -0.53 , whereas the least active compound has a partial charge of 0.260 . Both partial charges were centered on the nitrogen atom of the guanidine moiety in the P2 position.

The first step of the mechanism of serine proteases begins with the formation of the complex between the ligand and the receptor.⁹ Subsequently, the nonbonding electron pair of the oxygen atom of Ser135 performs a nucleophilic attack to an electrophilic carbon, which in this case involves the substrate aldehyde, and alkylates the enzyme.^{6,12} Docking studies can perform only the first step of this reaction (i.e., the formation of the complex). Thus, the D1 ligand has a shorter distance between the carbon atom of the formyl group (i.e., C–O Ser135; Table 1, 3.4 \AA), which is directly associated with the second step of the serine protease reaction. The investigated compounds were able to form a complex with the receptor with distances ranging between 8.63 and 8.87 \AA (**12m**). These results are in agreement with the binding energy, in which the most active compounds (i.e., **6m** and **12m**) are closer to the oxygen atom of Ser135, whereas the electrophilic carbons of the least active compounds (**1m** and **8m**) are farther from the reactive center.

The docking and QM/MM studies indicated that the complexes formed by ligands **6m** and **12m** possess lower binding affinities that are closer to the prototype inhibitor tetrapeptide D1. These results can be associated with their structures, the Ser135-carbonyl carbon ligand distances, steric effects and stabilization by water molecules. In general, the energies obtained by the QM/MM and molecular docking methodologies²¹ were compatible for some compounds, such as **6m** and **12m**, which strengthens their potential use as inhibitors of NS2B-NS3pro DENV-2. In other words, the QM/MM results corroborated the docking studies for compounds **6m** and **12m**. However, complementary studies using free energy and molecular dynamics simulations would be valuable to establish the utility of the present results.^{30,35}

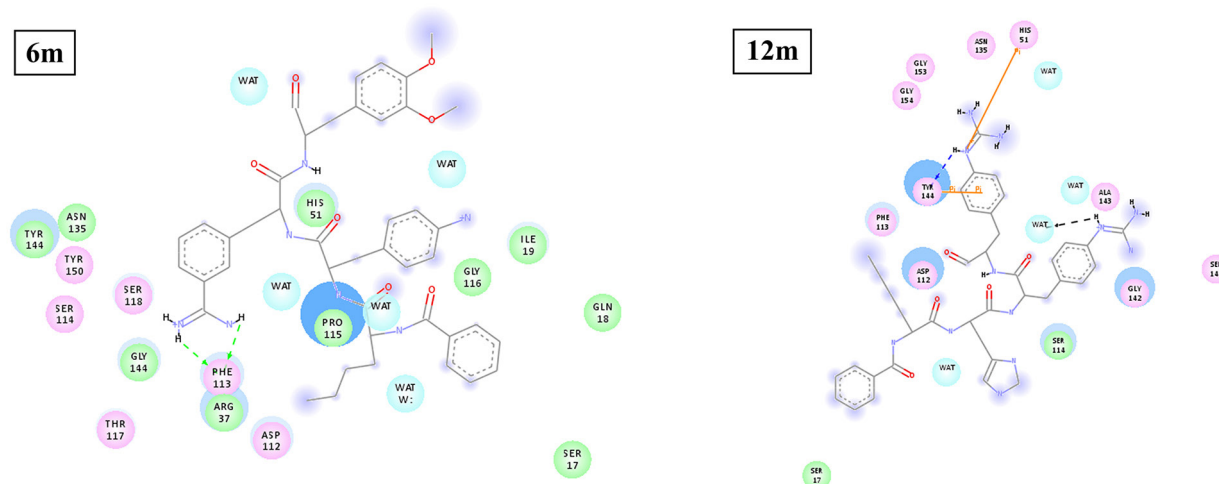


Figure 8. Interactions of compounds **6m** and **12m** in the active site of complex NS2B-NS3pro DENV-2, showing hydrogen bonds (pink circles with traced lines), hydrophobic interactions (green circles), pi-interactions (orange lines), water molecules (blue circles) and close contact residues (pink circles only). The solvent accessible regions are indicated as purple circles (in the molecule) and a blue halo (in the amino acid residue). This figure was construed using Discovery Studio Visualizer 3.1.⁴⁸

Conclusions

Of all compounds examined, ligands **6m** and **12m** demonstrated the lowest binding affinities by the molecular docking studies and QM/MM calculations. These ligands displayed the shortest distances between the carbon atom of the substrate aldehyde and the serine residue, which is essential for the catalytic mechanism. An analysis of the interactions observed for **6m** and **12m** relative to D1 showed a significant number of common interactions between these compounds and residues of the NS2B-NS3pro DENV-2 model. Specifically, Ser135, Gly151 and Tyr161 are amino acids that recognize inhibitors and direct them to this receptor. Thus, the participation of these residues is crucial for biological activity, especially because there is evidence in the literature to suggest that these amino acids contribute to the inhibition of DENV. In general, the results from molecular docking studies indicated that ligands with longer C–O Ser135 distances possessed higher binding affinities, which highlights the catalytic importance of the proximity of the carbon atom in the formyl group (i.e., aldehyde) to the oxygen atom of the serine residue. Clearly, each investigated compound possesses distinct intrinsic characteristics, and the present approach may impact the various ways in which each analysis is performed, such as designing its chemical structure, the number and type of intermolecular interactions and steric effects that contribute to molecular recognition. These findings may motivate virtual screening based on the pharmacophoric hypothesis to design new lead compounds and develop successful hits in the development of therapies against DENV, which is considered to be a neglected disease because there is no licensed treatment.

Supplementary Information

Supplementary material is available free of charge at <http://jbcs.sbq.org.br> as PDF file.

Acknowledgments

UFSJ/PPGBiotec, CNPq (449984/2014-1) and FAPEMIG (APQ-00557-14, CDS-APQ-02337-12, CBB-APQ-01301-11) are gratefully acknowledged for financial support, and I. P. G. is grateful to FAPEMIG for a fellowship.

References

- Bhatt, S.; Gething, P. W.; Brady, O. J.; *Nature* **2013**, *25*, 504.
- Simmons, C. P.; Farrar, J. J.; Chau, N. V.; Wills, B.; *N. Engl. J. Med.* **2012**, *366*, 1423.
- Tomlinson, S. M.; Malmstrom, R. D.; Watowich, S. J.; *Infect. Disord.: Drug Targets* **2009**, *9*, 1.
- Forshey, B. M.; Stoddard, S. T.; Halsey, E. S.; *Future Virol.* **2013**, *8*, 1145.
- Bollati, M.; Alvarez, K.; Assenberg, R.; Barontid, C.; Canard, B.; Cooke, S.; Coutard, B.; Decroly, E.; Lamballeried, X.; Gould, E. A.; Grard, G.; Grimes, J. M.; Hilgenfeld, R.; Jansson, A. M.; Malet, H.; Mancini, E. J.; Mastrangelo, E.; Mattevi, A.; Milani, M.; Moureaud, G.; Neyts, J.; Owens, R. J.; Renc, J.; Selisko, B.; Speronij, S.; Steuberg, H.; Stuart, D. I.; Unge, T.; Bolognesia, M.; *Antiviral Res.* **2010**, *87*, 125.
- Erbel, P.; Schiering, N.; D'Arcy, A.; Renatus, M.; Kroemer, M.; Lim, S. P.; Yin, Z.; Keller, T. H.; Vasudevan, S. G.; Hommel, U.; *Nat. Struct. Mol. Biol.* **2006**, *13*, 372.

7. Kuhn, R. J.; Zhang, W.; Rossmann, M. G.; Pletnev, S. V.; Corver, J.; Lenches, E.; Jones, C. T.; Mukhopadhyay, S.; Chipman, P. R.; Strauss, E. G.; Baker, T. S.; *Cell* **2002**, *108*, 717.
8. Niyomrattanakit, P.; Winoyanuwattikun, P.; Chanprapaph, S.; Angsuthanasombat, C.; Panyim, S.; Katzenmeier, G.; *J. Virol.* **2004**, *78*, 13708.
9. Nemukhin, A. V.; Grigorenko, B. L.; Rogov, A. V.; Topol, I. A.; Burt, S. K.; *Theor. Chem. Acc.* **2004**, *111*, 36.
10. Natarajan, S.; *Genet. Mol. Biol.* **2010**, *33*, 214.
11. Freceer, V.; Miertus, S.; *J. Comput.-Aided Mol. Des.* **2010**, *24*, 195.
12. Aleshin, A. E.; Shiryayev, S. A.; Strongin, A. Y.; Liddington, R. C.; *Protein Sci.* **2007**, *16*, 795.
13. Noble, C. G.; Seh, C. C.; Chao, A. T.; Shi, P. Y.; *J. Virol.* **2012**, *86*, 438.
14. Yin, Z.; Patel, S. J.; Wang, W.-L.; Wang, G.; Chan, W.-L.; Rao, K. R. R.; Alam, J.; Jeyaraj, D. A.; Ngew, X.; Patel, V.; Beer, D.; Lim, S. P.; Vasudevan, Subhash G.; Keller, T. H.; *Bioorg. Med. Chem. Lett.* **2006**, *16*, 36.
15. Yin, Z.; Patel, S. J.; Wang, W.-L.; Chan, W.-L.; Rao, K. R. R.; Wang, G.; Ngew, X.; Patel, V.; Beer, D.; Knox, J. E.; Ma, N. L.; Ehrhardt, C.; Lim, S. P.; Vasudevana, S. G.; Kellera, T. H.; *Bioorg. Med. Chem. Lett.* **2006**, *16*, 40.
16. Noble, C. G.; Shi, P.-Y.; *Antiviral Res.* **2012**, *96*, 115.
17. Leung, D.; Schroder, K.; White, H.; Fang, N.-X.; Stoermer, M. J.; Abbenante, G.; Martin, J. L.; Young, P. R.; Fairlie, D. P.; *J. Biol. Chem.* **2001**, *276*, 45762.
18. Godói, I. P.; Lima, W. G.; Alves, R. J.; Comar-Junior, M.; Ferreira, J. M. S.; Mizuno, C.; Taranto, A. G.; *Biochem. Biotechnol. Rep.* **2013**, *2*, 48.
19. Bordoli, L.; Kiefer, F.; Arnold, K.; Benkert, P.; Battey, J.; Schwede, T.; *Nat. Proc.* **2009**, *4*, 1.
20. Breda, A.; Basso, L. A.; Santos, D. S.; Azevedo Junior, W. F.; *Curr. Comput.-Aided Drug Des.* **2008**, *4*, 265.
21. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J.; *Nat. Rev. Drug Discovery* **2004**, *3*, 935.
22. Trott, O.; Olson, A. J.; *J. Comput. Chem.* **2009**, *31*, 455.
23. Vaqué, M.; Ardévol, A.; Bladé, C.; Salvadó, M. J.; Blay, M.; Fernández-Larrea, J.; Arola, L.; Pujadas, G.; *Curr. Pharm. Anal.* **2008**, *4*, 1.
24. Zhao, H.; Guo, Z.; *Drug Discovery Today* **2009**, *14*, 516.
25. Jaghoori, M. M.; Bleijlevens, B.; Olabarriaga, S. D.; *J. Comput.-Aided Mol. Des.* **2016**, *30*, 49.
26. Lima, W. E. A.; Pereira, A. F.; Castro, A. A.; Cunha, E. F. F.; Ramalho, T. C.; *Lett. Drug Des. Discovery* **2016**, *13*, 360.
27. Yotmanee, P.; Rungrotmongkol, T.; Wichapongb, K.; Choie, S. B.; Wahabe, H. A.; Kungwang, N.; Hannongbua, S.; *J. Mol. Graphics Modell.* **2015**, *60*, 24.
28. Taranto, A. G.; Carvalho, P.; Avery, M. A.; *J. Mol. Graphics Modell.* **2008**, *27*, 275.
29. Andrade, D. V. G.; Goes-Neto, A.; Comar-Junior, M.; Taranto, A. G.; *Int. J. Quantum Chem.* **2012**, *112*, 3164.
30. Lin, H.; Truhlar, D. G.; *Theor. Chem. Acc.* **2007**, *117*, 185.
31. Field, M. J.; Bash, P. A.; Karplus, M.; *J. Comput. Chem.* **1990**, *11*, 700.
32. Thompson, J. D.; Higgins, D. G.; Gibson, T. J.; *Nucleic Acids Res.* **1994**, *22*, 4673.
33. Breda, A.; Basso, L. A.; Santos, D. S.; Azevedo-Junior, W. F.; *Curr. Comput.-Aided Drug Des.* **2008**, *4*, 265.
34. Carregal, A. P.; Taranto, A. G.; Comar-Junior, M.; Lima, L. A.; Siqueira, J. M.; *Rev. Eletr. Farm.* **2011**, *8*, 71.
35. Sousa, S. F.; Fernandes, P. A.; Ramos, M. J.; *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 15.
36. Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R.; *J. Med. Chem.* **2005**, *48*, 6504.
37. Jones, G.; Willett, P.; Glen, R. C.; *J. Mol. Biol.* **1995**, *245*, 43.
38. Johansson, M. U.; Zoete, V.; Michielin, O.; Guex, N.; *BMC Bioinf.* **2012**, *13*, 1.
39. Warshel, A.; Levitt, M.; *J. Mol. Biol.* **1976**, *103*, 227.
40. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G.; *J. Phys. Chem.* **1996**, *100*, 19824.
41. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G.; *Chem. Phys. Lett.* **1995**, *246*, 122.
42. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P.; *J. Comput. Chem.* **2003**, *24*, 1999.
43. Tsui, V.; Case, D. A.; *Biopolymers* **2001**, *56*, 275.
44. Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A.; *Amber Version 11*, University of California, San Francisco, CA, USA, 2010.
45. Salomon-Ferrer, R.; Case, D. A.; Walker, R. C.; *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 198.
46. Fonseca, A. L.; Nunes, R. R.; Braga, V. M. L.; Comar-Junior, M.; Alves, J. R.; Varotti, F. P.; Taranto, A. G.; *J. Mol. Graphics Modell.* **2016**, *66*, 174.
47. Laskowski, R. A.; *J. Appl. Crystallogr.* **1993**, *26*, 283.
48. Accelrys Software Solutions Pvt. Ltd.; *Discovery Studio Visualizer Version 3.1*; Accelrys Inc., San Diego, CA, USA, 2011.
49. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.;

- Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; *Gaussian Version 09*, Gaussian Inc., Wallingford, CT, USA, 2009.
50. Lau, Y. K.; Nishizawa, K.; Tse, A.; Brown, R. S.; Kebarle, P.; *J. Am. Chem. Soc.* **1981**, *103*, 6291.
51. Tripathi, S. K.; Soundarya, R. N.; Singh, P.; Singh, S. K.; *Chem. Biol. Drug. Des.* **2015**, *85*, 18.
52. Santos-Junior, M. C.; Assis, S. A.; Góes-Neto, A.; Duarte, A. A.; Alves, R. J.; Comar-Junior, M.; Taranto, A. G.; *Chem. Cent. J.* **2013**, *7*, 1.
53. Groenhof, G. In *Biomolecular Simulations: Methods and Protocols, Methods in Molecular Biology*; Monticelli, L.; Salonen, E., eds.; Springer Science: New York, USA, 2013, ch. 3.
54. Kleinjung, J.; Fraternali, F.; *Curr. Opin. Struct. Biol.* **2014**, *25*, 126.
55. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T.; *J. Am. Chem. Soc.* **1990**, *112*, 6127.
56. Eisenberg, D.; Lüthy, R.; Bowie, J. U.; *Methods Enzymol.* **1997**, *277*, 396.
57. Thilagavathi, R.; Mancera, R. L.; *J. Chem. Inf. Model.* **2010**, *50*, 415.
58. Stewart, J. J. P.; *J. Mol. Model.* **2007**, *13*, 1173.
59. Weiner, S. J.; Seibel, G. L.; Kollman, P. A.; *Biophysics* **1986**, *83*, 649.
60. Giese, T. J.; York, D. M.; *J. Chem. Phys.* **2007**, *127*, 1.
61. Dubey, K. D.; Chaubey, A. K.; Ojha, R. P.; *Med. Chem. Res.* **2012**, *21*, 1030.
62. Rossm, J. M. V.; *J. Pharm. Pharmacol.* **2011**, *15*, 285.

Submitted: April 13, 2016

Published online: August 22, 2016