Interaction of ZIKV NS5 and STAT2 Explored by Molecular Modeling, Docking, and Simulations Studies

Gerardo Armijos-Capa[1,2], Paúl Pozo-Guerrón[1,2], F. Javier Torres[1,2], and Miguel M. Méndez[1,2(✉)]

[1] Instituto de Simulación Computacional (ISC-USFQ), Universidad San Francisco de Quito, Diego de Robles y Vía Interoceánica, Quito 17-1200-841, Ecuador
mmendez@usfq.edu.ec

[2] Departamento de Ingeniería Química, Grupo de Química Computacional y Teórica (QCT-USFQ), Universidad San Francisco de Quito, Diego de Robles y Vía Interoceánica, Quito 17-1200-841, Ecuador

Abstract. ZIKV NS5 has been associated with inhibition of type I IFN during the host antiviral response. The protein-protein interaction may promote the proteasomal degradation of STAT2, although the entire mechanism is still unknown. In this study, a three-dimensional model of the full STAT2 protein (C-score = −0.62) was validated. Likewise, the top scored docked complex NS5-STAT2 is presented among several other models; the top model shows a total stabilizing energy for the complex of $-77.942\,\mathrm{kcal\,mol^{-1}}$ and a Gibbs binding free energy of $-4.30\,\mathrm{kcal\,mol^{-1}}$. The analysis of the complex has revealed that the interaction is limited to three domains known as N-terminal from STAT2 and Mtase-Thumb from NS5; both located in the ordered regions of these proteins. Key residues involved in the interaction interface that showed the highest frequency among the models are stabilized by electrostatic interactions, hydrophobic interactions, salt bridges, and ionic interactions. Therefore, our findings support the experimental preliminaries observations reported in the literature and present additional structural characterization that will help in the drug design efforts against ZIKV NS5.

Keywords: Zika virus · NS5 · STAT2 · Molecular dynamics · Docking

1 Introduction

Since 2015, the outbreak of Zika virus (ZIKV) in Central and South America has become a worldwide health concern. ZIKV infection is associated with congenital diseases such as microcephaly and rare but severe complications in adults such as

Supported by Universidad San Francisco de Quito.

I. Rojas et al. (Eds.): IWBBIO 2019, LNBI 11466, pp. 165–176, 2019.
https://doi.org/10.1007/978-3-030-17935-9_16

the Guillain-Barré syndrome [1,2]. Zika virus is a mosquito-borne flavivirus that has a genome of single-strand positive RNA of 10 kb that encodes ten proteins namely the envelope (E), membrane precursor (PrM), and capsid (C) which contribute to the viral particles. It also encodes for seven nonstructural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) which contribute to viral replication [2]. Among NS proteins, NS5 is the largest Zika protein with a weight of ∼103 kDa and consists in two principal domains. The first one is known as the methyltransferase (Mtase) which is related to the decrease of host innate immune response and promotes the translation of the polyprotein through the addition of 5' RNA cap structure. The second one is the RNA-dependent RNA polymerase (RdRp) required for the initiation and elongation of RNA synthesis [3]. The viral response at human cells is set up by intracellular pattern recognition receptors (PRRs) proteins. They recognize viral pathogen marks such as viral RNA, DNA or protein in the type I interferon (IFN) signaling. This recognition induces activation of innate immune signaling, leading to the up-regulation of IFN-stimulated genes (ISGs). Flaviviruses such as West Nile virus (WNY), Dengue virus (DENV) or Yellow Fever virus (YFV) use various strategies of IFN antagonism [4]. However, flaviviruses share replication strategies based on the formation of a polyprotein which can inhibit transcriptional activation of IFNs and ISGs during virus infection [5–7]. In this vein, NS5 is considered as a potent and specific antagonist of type I IFN signaling [4]. ZIKV NS5 has been associated with inhibition of type I IFN during the host antiviral response, because it may promote the proteasomal degradation of STAT2. This role has been substantiated by different experimental studies which have determined that strains of ZIKV antagonize type I IFN where NS5 reduces the STAT2 level and prevents the translocation of STAT2 from the cytoplasm to nuclei in immunoprecipitation assays in 293 T cells [4]. Consequently, there is inhibition of the IFN induction of the ISG [4]. In others test, STAT2 levels have been compared in cells expressing each of the two functional domains of NS5. The comparison has shown that with only the Mtase domain expressed, the levels of STAT2 degradation were higher. While with only RdRp domain expressed, the levels of STAT2 degradation were negligible [7]. On the other hand, the Mtase activity in full-length NS5 has not shown to be necessary for the STAT2 degradation, suggesting that other regions of the protein NS5 may contribute to degradation mechanism [7]. The interaction with other proteins in the pathway of type I IFN were reviewed through other essays. It has shown that reduced STAT2 levels associated to ZIKV infection are independent of the presence of proteins such as ubiquitin ligase UBR4 or ubiquitin-specific protease USP 18; hence the interaction between STAT2-NS5 is directly related with degradation of STAT2, although the complete mechanisms is still unknown [4–7]. The combination of different computational tools such as molecular modeling, molecular docking, and molecular dynamic (MD) simulation permit to study a binary complex contributing to provide a clear protein-protein binding mechanism that could lead to significant advances in the understanding of NS5-STAT2 complex. Therefore, the aim of the present study is to elucidate the potential role and mechanism of the NS5-STAT2

interaction involved in pathway of type I IFN signaling during ZIKV infection from a computational approach.

2 Methods

Molecular modeling and quality check. The NS5 three-dimensional model (PDB code: 5TMH) was obtained from the protein data bank. However, only two fragments (PDB codes: 5OEN and 2KA4) have been reported for STAT2. Three-dimension models were generated through algorithms of protein threading and homology alignment programs known as I-TASSER and Phyre2. In order to evaluate the NS5 and STAT2 models quality the programs Verify-3D, ERRAT, and PROCHECK were used. STAT2 experimental fragments reported in the database were employed to perform a structural alignment with the STAT2's model through TM-align. **MD simulation and clustering.** GROMACS 5.1.2 [8] was used to perform the MD simulation of NS5 and STAT2 as well as NS5-STAT2 complex. MD simulations were conducted by using AMBER 03 force field. The systems were established using the following features; cubic boxes filled with SPC216 water molecules, TIP3P water, Na^+ and Cl^- counterions were added to neutralize the system with a concentration of 0.1 M in order to simulate physiological conditions of the cells and periodic boundary conditions. PME was used for non-bonded interactions such as electrostatic interaction and van der Waals with a cut-off of 12 Å and a 2 fs time step during the simulation. The energy minimization was obtained through the steepest-descent algorithm, and the maximum force of the system was set to 100 kJ $(\text{mol nm})^{-1}$ on all atoms. NVT and NPT ensembles were equilibrated using Berendsen thermostat and Nose-Hoover thermostat for 500 ps. Parrinello-Rahman barostat was employed to maintain the pressure isotropically with a value of 1.0 bars and compressibility of $4.5 \times 10^{-5} bar^{-1}$. The systems were submitted to 50 ns of production which was initialized using output data retrieved from previously run equilibration simulation at 310 K and 1 atm. Besides, all bonds length containing hydrogen were constrained using the Linear Constraint Solver (LINCS) algorithm. Gromacs utilities was used to analyze MD trajectory, and the charts were plotted using Grace. In order to explore the structural conformations generated from MD trajectories, a clustering was made. The clustering is based on Root Mean Square Distance (RMSD) of Cα atoms with a cut-off of 2.2 Å for each trajectory through the GROMOS clustering algorithm which is executed in the *gmx cluster* tool of GROMACS 5.1.2. A representative structure of each protein's trajectory was extracted in order to be used in an ensemble docking. **Docking and Gibbs binding free energy.** Docking complex was performed using ClusPro 2.0 and Pydock. Each binary complex selected was quantified by the binding strength of the protein-protein interface, Gibbs binding free energy (ΔG_{Bind}) and interface area through the programs PPCheck, FoldX and PISA, respectively. **Protein-protein interaction and electrostatic calculation.** Protein Interaction Calculator (PIC) web server and COCOMAPS were used to analyzed the protein-protein interaction between binary complexes. The lowest energy structure of

NS5, STAT2, and NS5-STAT2 complexes were used to perform the electrostatic calculations through the Poisson-Boltzmann (PB) equation implemented in the APBS Program as plug-in added in the program PyMOL 2.2.2.

3 Results and Discussion

3.1 Molecular Modeling and Validation of STAT2

The three-dimensional model of STAT2 was generated using the amino acid sequence by two different approaches namely protein threading of I-TASSER and homology of Phyre2. Through, I-TASSER five models were obtained. The best model has been chosen according to the highest C-score whose value of −0.62 suggests a correct topology. Likewise, the best model of Phyre2 has been obtained according to the identity percentage and confidence values which were 41% and 100%, respectively. This model was built with a template of the non-phosphorylated STAT1 (PDB code: C1YVIB). Both the best model of I-TASSER and Phyre2 were compared by structural quality to select the model to be used in the next analyses. The results have shown that STAT2's model of I-TASSER obtains a quality factor of 70.74% better than Phyre2's model with a value of 56.65% according to Verify-3D. With ERRAT the trend was similar, because I-TASSER's model presents better value than Phyre2's model, for this latter model an unacceptable value of 21.69% unlike to the value of 77.80% achieved by I-TASSER's model. Ramachandran plots were obtained by PROCHECK where Phyre2's model has reached more appropriated values for favored and allowed regions than I-TASSER's model since gathers values of 75% and 17.2%, respectively. These latter values are higher than the I-TASSER's model with values of 72.6% and 21.6%. However, Phyre2's model has gathered a higher percentage in the disallowed regions indicating that the stereochemical quality is inferior to I-TASSER'S model. Based on these outcomes, we selected the I-TASSER's model because their residues are positioned in the favorable three-dimensional structure environment in agreement with the experimental functional domains as shown in Fig. 1. Modeling by homology is usually more accuracy wherein stereo-chemical restrains and segments matching are then considered. However, it is a limited method because a template must have a high identity with target sequence. Generally, the accuracy of homology models under to 30% is due to alignment error leading to incapability to generate structures that fit with the target sequences [9,10]. Otherwise, modeling by threading is capable to build a model without a template. It represents an improvement when the target sequence is unknown or has segments (surface loops) that are misaligned with template sequences. Therefore, this latter is considered more successful in contrast to an approach by homology [9,10]. Additionally, STAT2's model has been compared with experimental fragments of STAT2 in order to observe the structural similarity. Each fragment (2KA4 and 5OEN) was aligned with STAT2's model where the fragment 50EN has been matched by 168 residues with a TM-score of 0.86 in the Coiled-coil domain while the fragment 2KA4 has been paired by 39 residues with a TM-score of 0.34 in the Transactivation domain.

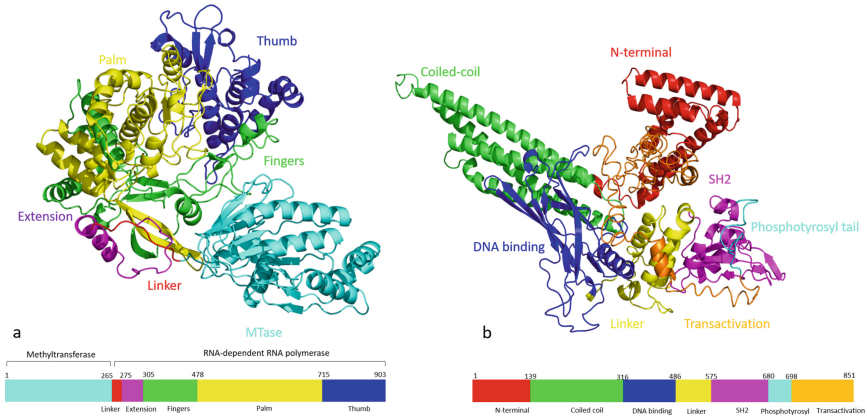

Fig. 1. Three-dimensional structures of (a) NS5 and (b) STAT2 model. The functional domains of each model are shown in different color. NS5: Mtase = Methyltransferase domain (1–264 aa), Linker domain (265–275 aa), Extension (275–304 aa), Fingers domain (305–477 aa), Palm domain (478–714 aa) and Thumb domain (715–903 aa) STAT2: N-terminal domain (1–138 aa) involved in dimerization/-tetramerization, Coiled-coil domain (139–315 aa) involved in interaction with other proteins, DNA binding domain (316–485 aa), Linker domain (486–574 aa), SH2 domain (575–679 aa), Phosphotyrosyl tail segment (680–697 aa), and Transactivation domain (698–851 aa). (Color figure online)

This latter one has hardly coincided with the STAT's model, because it is in a disorder region with a coiled-coil as secondary structure. Nevertheless, the STAT2's model obtained is suitable to be used as a reference protein model of STAT2.

3.2 Analysis of Docking Complex NS5-STAT2

In order to understand the biological functions as well as interactions of these proteins, it is essential to explore the conformational changes in short time lapses at the atomic level. MD simulation is a tool that describes atomic motion, structural properties and thermodynamic behavior of the systems at equilibrium [11,12]. Besides, MD simulation is employed to improve the virtual screening in the docking processes since it can be considered a step of refining such models. This refining provides flexibility and structural rearrangements to proteins allowing to obtain in overall a more realistic structure of each model since the methods of protein-protein docking are usually supported by the shape complementarity [13–16]. Therefore, NS5's x-ray structure and STAT2's model were subjected to 50 ns MD simulation protocol defined in the method section to minimize and equilibrate at near physiological conditions. From the trajectories of NS5 and STAT2, it was performed a clustering analysis to explore the conformation of proteins generated by MD simulation. According to the RMSD cut-off detailed in the methods section, the dominant clusters for NS5 and STAT2 constitute

∼75% of total protein structures. Representative structures were extracted from these cluster as targets to build an ensemble docking [13]. Each structure has the largest stabilizing energy which is considered close to a native structure [17]. In order to explore the interaction between these proteins, docking analyses have been performed through ClusPro and Pydock. At this stage, two troubles arise to predict the correct solution, the first one is the scoring functions of ClusPro and Pydock and second one is the binding site which is unknown [14]. Although some algorithms are able to score a list of possible structures, however they are not reliable to discriminate false positives, that is, complexes with a high score but with a low rank [14]. Therefore, in order to classify docked complexes we used the software PPCheck with a process of discrimination by functional domains, and ΔG_{Bind}. PPCheck quantifies the strength of protein-protein interaction using pseudo-energies. The outcomes have shown that the total stabilizing energy for Pydock's models is positive ($\Delta E > 0$) while that for ClusPro's models are negative ($\Delta E < 0$). The total stabilizing energy values that have negative energy tendency are related with the increase of the number of interface residues [18]. Hence, negative energy suggests that there is a contact interface between proteins. In contrast, the interaction is deficient when the contact areas are almost null in systems with positive energy. Among of them, the ClusPro's models 4, 7 and 9 have obtained the largest total stabilizing energy with a normalized energy per residue closed to -6kJ mol^{-1}. Therefore, they have fallen in a correct docking pose [18]. Cluspro's models have also shown to interact via known domains in both proteins. A test of total stabilizing energy and normalized energy per residue among the N-terminal domain from STAT2 and Mtase and Thumb domains from NS5 have been performed. The models 1, 4, and 7 have demonstrated that the interaction between these domains has gathered the largest total stabilizing energy. However, the interaction in the model 9 is only between the N-terminal and Mtase domains and presents a negligible total stabilizing energy (see Table 1). Likewise, the models 1, 4 and 7 have reached the largest normalized energy per residue close to -6kJ mol^{-1} showing that these models involve a correct docked position [18]. In contrast to the other models that do not show interaction among these domains. In order to understand the difference between the model of Cluspro and Pydock was used the surface area. It provides features in the protein-protein interaction which displays a high degree of structural complementarity and chemical complementarities [19]. As a result, the interface areas of Cluspro's models ($\overline{x} = 1973.59\,\text{Å}^2$) are larger than any Pydock's model ($\overline{x} = 1068.03\,\text{Å}^2$). Similar results to our ClusPro's models have been found in a published set of 75 crystal structures that present an interface area average of $2000\,\text{Å}^2$ in each member. This average is considered a specific protein-protein interaction with high complementarity [19,20]. Likewise, It has been observed that protein-protein interactions have large contact surfaces ($1500–3000\,\text{Å}^2$). In contrast, the contact area between small molecules and proteins targets has been estimated between 300 to $1000\,\text{Å}^2$ [21]. The ΔG_{Bind} has been calculated using empirical force field of FoldX for the atomic coordinates of Cluspro's models [22]. In this initial analysis, each model has achieved a positive

value of interaction energy $\Delta G_{Bind} > 0$. It means that the docked complexes may be considered as unstable structures since it denotes energy-unfavorable coupling between both proteins. This unstable state between proteins NS5-STAT2 can also be considered as a non-covalent interaction with negative cooperativity since the affinities between ligand-receptor are decreased. Then, the docked complex is less well bonded and their atoms exhibit major internal motions [23]. A negative cooperativity also suggests that the docked complexes will need a great amount of energy (highly endothermic) to couple because the binding between NS5 and STAT2 is not a process spontaneous. Therefore, $\Delta G_{Bind} > 0$ is related with a favorable change entropy [23]. Additionally, the non-covalent interaction with a negative cooperativity in the interface between NS5-STAT2 may be influenced by coupling sites geometry. A coupling of non-covalent interaction with positive cooperative causing a structural loosening in the docked complex because the contact distance between ligand-receptor is greater [23]. The models 1, 7, 4, and 9 show the largest binding energy (ΔG_{Bind}) in comparison with the remaining models. However, we decided to take into consideration only models 1, 4, and 7 to be subjected to MD simulation because they showed interaction in the Mtase and Thumb domains with the N-terminal domains in the NS5-STAT2 complex.

Table 1. Results of total stabilizing energy and normalized energy per residue for ten ClusPro's models based on the functional domains of NS5 and STAT2. The functional domains involved in the interaction of complex were the N-terminal from STAT2 and Mtase and Thumb from NS5.

Models	Total stabilizing energy (kcal mol^{-1})		Normalized energy per residue (kJ mol^{-1})	
	N-Ter/Mtase	N-Ter/Thumb	N-Ter/Mtase	N-Ter/Thumb
Model 1	−23.2	−27.192	−2.7	−1.9
Model 2	-	-	-	-
Model 3	-	-	-	-
Model 4	−19.116	−48.282	−3.33	−3.61
Model 5	-	-	-	-
Model 6	−9.037	-	−1.18	-
Model 7	−59.388	−39.959	−3.5	−3.89
Model 8	−3.774	-	−0.88	-
Model 9	−1.114	-	−4.66	-
Model 10	-	-	-	-

3.3 Contact Map and Electrostatic Analysis

In order to confirm the interaction among Mtase and Thumb with N-terminal domains, a contact map of intramolecular interactions of the NS5-STAT2 complex is illustrated in Fig. 2. The distance range between two proteins has been marked with red, yellow, green and blue for 7 Å, 10 Å, 13 Å, and 16 Å, respectively. Largest regions of two partners that are in contact have been located in

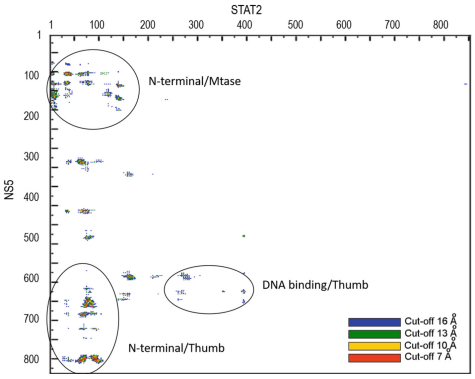

Fig. 2. Contact map of NS5 and STAT2 shows the intermolecular contacts at reducing distances which are red = 7 Å, yellow = 10 Å, green = 13 Å and blue = 16 Å. The circles identify the domains related such as N-terminal/Mtase, N-terminal/Thumb and Thumb/DNA binding. (Color figure online)

the map. The regions associated corresponds to first residues of Mtase domain (1–200) and the last residues of Thumb domain (650–750) in the NS5 while that for region of STAT2 is in the residues of N-terminal domain (1–150) and a smaller interaction region in the residues from 300 to 400 that correspond to DNA binding domain. The contact map confirms our previous results since the protein-protein interaction is mainly focused on in the domains described. An electrostatic potential analysis has been performed because electrostatic interactions are favored for the protein-protein interaction as well as the stabilization in a complex. The electrostatic role depends on the type of hetero- or homocomplexes, that is, a complex is formed by different or identical proteins which have net charge (positive or negative). In our case, the complex NS5-STAT2 is a heterocomplex which carries an opposite net charge that leads to the attraction between proteins. However, the arrangement of heterocomplex will be limited by residues distribution that will change global net charge of each protein at short distances in the interface [24]. Figure 3 shows the NS5-STAT2 complex. In the region of the N-terminal domain of STAT2, their buried residues are highly polar. In contrast, the residues located in the cavity of NS5 have a hydrophobic character with few polar groups around. Studies suggest that polar residues in the interface favorably contribute in two ways; the first, through a specific association between proteins, and second improving the stabilization of complexes since the interacting forces do not need to be strong for the formation of complexes. Moreover, regions in the protein with polar groups are usually sites known as hot spots which are crucial for a better affinity between proteins [24].

3.4 MD Simulation of NS5-STAT2 Complexes

The models 1, 4, and 7 have been minimized to optimize the complex geometry and check the Gibbs binding free energy (ΔG_{Bind}) again. ΔG_{Bind} has been

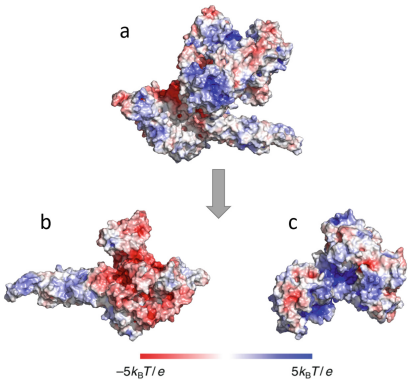

Fig. 3. A view of electrostatic surface potential for (a) NS5-STAT2 complex, (b) STAT2 and (c) NS5. Red is negative charge, blue is positive charge, and white is neutral. (Color figure online)

calculated for representative conformations (snapshots) of each trajectory in the models 1, 4, and 7, in order to verify the behavior of the complex NS5-STAT2. According to the clustering of MD trajectories, each model has obtained 6001 structures which have been clustered in 33, 22, and 21 groups for models 1, 4, and 7, respectively [25]. Hence, the results have shown that ΔG_{Bind} of model 1 and 7 ($\Delta G_{Bind} < 0$) are $-4.30\,\mathrm{kcal\,mol^{-1}}$ and $-1.67\,\mathrm{kcal\,mol^{-1}}$, respectively while for model 4 ($\Delta G_{Bind} > 0$) is $0.27\,\mathrm{kcal\,mol^{-1}}$. The coupling between the proteins is energetically favorable in models 1 and 7 in opposition to model 4. The same range of values of Gibbs binding free energy $\Delta G_{Bind} < 0$ are also obtained in experimental measurements of binding affinity for protein-protein complexes tested on a benchmark of 144 structures [26]. Likewise, a study with an empirical approach to calculate the Gibbs binding free energy based on three variables of the interface in complexes has estimated values almost close to model 1 [27]. These latter results have demonstrated that applying MD simulation to the system is possible to improve the values of Gibbs binding free energy. On the other hand, the initial atomic coordinates of models have shown to reach a $\Delta G_{Bind} > 0$ meaning that docked complexes has non-covalent interactions with negative cooperativity. However, after a MD simulation, the models 1 and 7 have gathered a $\Delta G_{Bind} < 0$ which benefit the non-covalent interactions with positive cooperativity [23]. This change in the cooperativity is acceptable in the same system. The motion of atoms of the protein by the MD simulation produces a new stable state that has a net effect in the thermodynamic parameters [12, 23]. Hence, non-covalent interaction with positive cooperativity is related to an exothermic binding which allows an increment in the bonding ligand-receptor. The improving in the binding is associated with a favorable enthalpy and adverse in entropy because with a strong coupling the internal motions of the complex is reduced [23].

3.5 Interaction NS5-STAT2

The protein-protein interaction is mediated by domain-domain interactions. In our analyses, it has been identified that the interaction is given by the N-terminal domain from STAT2 and Mtase and Thumb domains from NS5 [28]. The interactions in the interface of NS5-STAT2 are stabilized by electrostatic interactions, hydrophobic interactions, salt bridges, and ionic interactions. In the case of NS5, 33 residues are in the contact area with an elevated frequency in all models. Moreover, 16 of them have presented one o more types of interactions. The NS5 residues involved are Arg-163, Arg-175, Arg-37, Arg-57, Arg-681, Arg-84, Arg-856, Glu-149, Leu-847, Lys-105, Lys-331, Pro-108, Pro-857, Trp-848, Val-335, and Val-336. These residues are located in the Mtase and Thumb domains. Regarding STAT2, 55 residues have a high frequency where 19 of them are involved with at least one type of interaction. The residues are Arg-796, Arg-88, Asp-77, Asp-794, Asp-850, Asp-93, Glu-40, Glu-715, Glu-722, Glu-79, Glu-801, Glu-804, Glu-814, His-85, Leu-684, Leu-691, Leu-727, Leu-81, and Lys-89. They have been located in the N-terminal domain. Studies of hot-spots in the binding sites of protein-protein interfaces have located a particular enrichment of Trp, Try and Arg, as well as a high presence of polar residues [29]. In contrast, hydrophobic residues as Val and Leu are associated to interfaces largely hydrophobic and nonpolar surface areas [29]. In our case, Arg and Glu are the hydrophilic residues with the largest presence in the interface of both NS5 and STAT2. Besides, hydrophilic residues as Lys, Pro, and Asp are also displayed in the interface of both proteins. Other studies have concluded that residues as Trp, Met, Try, Phe, Cys, and Ile are frequently in the binding interfaces. Residues as Tyr, Trp, His, and Cys have been traced in high-affinity interfaces in comparison with low-affinity interfaces [30]. His and Trp are present in the interface of NS5-STAT2 complex. Hence the interface shows a high-affinity. However, residue as Lys is located in low-affinity interface. In our outcomes, this last residue is found in the interface of NS5-STAT2 complex which will counteract a possible high-affinity in the NS5-STAT2 complex [30]. As was mentioned before, NS5-STAT2 is a heterocomplex which has differences properties associated with amino acid composition, contact sites, and interface area [31]. Hence, studies of homo- and heterocomplexes have revealed that the amino acid interface composition is different between them. The residues in heterocomplex have a greater prevalence and propensity to be in the interface are Leu, Val, Ile, Arg, Tyr, Trp, Met, and Phe [31]. These types of residues are also prevalent in our findings except for Ile, Met, and Phe.

4 Conclusions

The computational approach has permitted to analyze the different structural and dynamics features of NS5, STAT2 and the interaction of NS5-STAT2 complex. The STAT2's structure from I-TASSER was the selected model according to quality analyses. Moreover, it displays a high correlation with experimental

fragments of STAT2. Three docked complexes provided by ClusPro showed interaction on three domains (N-terminal/Mtase and N-terminal/Thumb) enriched with polar residues. MD simulations on the three docked complexes benefits the interaction between proteins because the behavior of proteins is affected by internal atomic motions. Thus, they has shown a $\Delta G_{Bind} < 0$ where the best docked complex have a ΔG_{Bind} of $-4.30\,\mathrm{kcal\,mol^{-1}}$. On the other hand, the NS5-STAT2 docked complex has revealed that the key interacting residues are stabilized by electrostatic interaction, hydrophobic interaction, salt bridges, and ionic interaction. Therefore, this study sheds light in the interaction of NS5-STAT2 as support of the experimental studies and the development of drugs against ZIKV NS5.

Acknowledgments. Our thankful to Universidad San Francisco de Quito for the use of the High Performance Computing System-USFQ.

References

1. Cox, B.D., Stanton, R.A., Schinazi, R.F.: Predicting Zika virus structural biology: challenges and opportunities for intervention. Antiviral Chem. Chemother. **24**(3–4), 118–126 (2015)
2. Wang, B., et al.: The structure of Zika virus NS5 reveals a conserved domain conformation. Nat. Commun. **8**, 14763 (2017)
3. Zhao, B., et al.: Structure and function of the Zika virus full-length NS5 protein. Nat. Commun. **8**, 1–9 (2017)
4. Grant, A., et al.: Zika virus targets human STAT2 to inhibit type I interferon signaling. Cell Host Microbe **19**(6), 882–890 (2016)
5. Arimoto, K., et al.: STAT2 is an essential adaptor in USP18-mediated suppression of type I interferon signaling. Nat. Struct. Mol. Biol. **24**(3), 279–289 (2017)
6. Bowen, J.R., et al.: Zika virus antagonizes type I interferon responses during infection of human dendritic cells. PLoS Pathog. **13**(2), e1006164 (2017)
7. Kumar, A., et al.: Zika virus inhibits type-I interferon production and downstream signaling. EMBO Rep. **17**(12), 487–524 (2016)
8. Abraham, M., Hess, B., van der Spoel, D., Lindahl, E.: GROMACS User Manual version 5.0.7 (2015). www.Gromacs.org
9. Fiser, A.: Template-based protein structure modeling. Methods Mol. Biol. **673**, 1–20 (2010)
10. Schwede, T., Sali, A., Eswar, N.: Protein structure modeling. In: Schwede, T., Peitsch, M.C. (eds.) Computational Structural Biology: Methods and Applications, chap. 1, pp. 1–33. World Scientific Publishing Co., Pte. Ltd., Danvers (2008)
11. Vlachakis, D., Bencurova, E., Papangelopoulos, N., Kossida, S.: Current state-of-the-art molecular dynamics methods and applications, 1 edn, vol. 94. Elsevier Inc. (2014)
12. Karplus, M., Petsko, G.A.: Molecular dynamics simulations in biology. Nature **347**(6294), 631–639 (1990)
13. Hospital, A., Goñi, J.R., Orozco, M., Gelpi, J.: Molecular dynamics simulations: advances and applications. Adv. Appl. Bioinf. Chem. **8**, 37–47 (2015)
14. Halperin, I., Ma, B., Wolfson, H., Nussinov, R.: Principles of docking: an overview of search algorithms and a guide to scoring functions. Proteins Struct. Funct. Genet. **47**(4), 409–443 (2002)

15. Brooijmans, N., Kuntz, I.D.: Molecular recognition and docking algorithms. Ann. Rev. Biophys. Biomol. Struct. **32**(1), 335–373 (2003)
16. Smith, G.R., Sternberg, M.J.: Prediction of protein-protein interactions by docking methods. Curr. Opin. Struct. Biol. **12**(1), 28–35 (2002)
17. Sikosek, T., Chan, H.S.: Biophysics of protein evolution and evolutionary protein biophysics. J. R. Soc. Interface **11**(100), 20140419–20140419 (2014)
18. Sukhwal, A., Sowdhamini, R.: Oligomerisation status and evolutionary conservation of interfaces of protein structural domain superfamilies. Mol. BioSyst. **9**(7), 1652–1661 (2013)
19. Elcock, A.H., Sept, D., Mccammon, J.A.: Computer simulation of protein protein interactions. J. Phys. Chem. B **105**, 1504–1518 (2001)
20. Keskin, O., Tuncbag, N., Gursoy, A.: Characterization and prediction of protein interfaces to infer protein-protein interaction networks. Curr. Pharm. Biotechnol. **9**(2), 67–76 (2008)
21. Smith, M.C., Gestwicki, J.E.: Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. Expert Rev. Mol. Med. **14**, 1–24 (2012)
22. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., Serrano, L.: The FoldX web server: an online force field. Nucleic Acids Res. **33**(Suppl. 2), 382–388 (2005)
23. Williams, D.H., Stephens, E., O'Brien, D.P., Zhou, M.: Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. Angew. Chem. Int. Ed. **43**(48), 6596–6616 (2004)
24. Zhang, Z., Witham, S., Alexov, E.: On the role of electrostatics on protein-protein interactions. Phys. Biol. **8**(3), 035001 (2011)
25. Snyder, P.W., Lockett, M.R., Moustakas, D.T., Whitesides, G.M.: Is it the shape of the cavity, or the shape of the water in the cavity? Eur. Phys. J.: Spec. Top. **223**(5), 853–891 (2014)
26. Vreven, T., Hwang, H., Pierce, B.G., Weng, Z.: Prediction of protein-protein binding free energies. Protein Sci. **21**(3), 396–404 (2012)
27. Ma, X.H., Wang, C.X., Li, C.H., Chen, W.Z.: A fast empirical approach to binding free energy calculations based on protein interface information. Protein Eng. **15**(8), 677–681 (2002)
28. Brito, A.F., Pinney, J.W.: Protein-protein interactions in virus-host systems. Front. Microbiol. **8**(Aug), 1–11 (2017)
29. Ma, B., Elkayam, T., Wolfson, H., Nussinov, R.: Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. Proc. Nat. Acad. Sci. **100**(10), 5772–5777 (2003)
30. Erijman, A., Rosenthal, E., Shifman, J.M.: How structure defines affinity in protein-protein interactions. PLoS ONE **9**(10), e110085 (2014)
31. Talavera, D., Robertson, D.L., Lovell, S.C.: Characterization of protein-protein interaction interfaces from a single species. PLoS ONE **6**(6), e21053 (2011)