## Supplementary Information

## Adjacent dimer epitope of envelope protein as an important region for Zika virus serum neutralization: a computational investigation

Authors:<br>Carlos Alessandro Fuzo ${ }^{1,2^{*}}$<br>E-mail: cafuzo@usp.br<br>Luiz Felipe Lemes de Araujo ${ }^{1,3}$<br>E-mail: araujolfelipe@usp.br<br>Rafael de Souza Pontes ${ }^{1,3}$<br>E-mail: rafaelpontes@usp.br<br>Patricia Martinez Évora ${ }^{1,3}$<br>E-mail: patievora@hotmail.com<br>Rodrigo Guerino Stabeli ${ }^{1,3}$<br>E-mail: rodrigo.stabeli@fiocruz.br

Affiliations:
${ }^{1}$ Plataforma de Pesquisa em Medicina Translacional, Fundação Oswaldo Cruz Fiocruz SP, Avenida dos Bandeirantes 3900, Campus USP Ribeirão Preto, São Paulo, Brazil.
${ }^{2}$ Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Avenida Bandeirantes 3900, Ribeirão Preto, SP, Brazil.
${ }^{3}$ Departamento de Bioquímica e Imunologia, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Avenida Bandeirantes 3900, Ribeirão Preto, São Paulo, Brazil.
*Corresponding Author
E-mail:cafuzo@usp.br


Figure S1 - Diagnostic plots for native system simulation showing the profiles of potential energy (A), root-mean-square deviation after the superposition of $\mathrm{C}_{\alpha}$ atoms to initial structure (B), gyration radius for all atoms (C) and root-mean square fluctuation of $\mathrm{C}_{\alpha}$ atoms from 100 to 200 ns (D).


Figure S2 - Diagnostic plots for Asp67Ala system simulation showing the profiles of potential energy (A), root-mean-square deviation after the superposition of $\mathrm{C}_{\alpha}$ atoms to initial structure (B), gyration radius for all atoms (C) and root-mean square fluctuation of $\mathrm{C}_{\alpha}$ atoms from 80 to 100 ns (D).


Figure S3 - Diagnostic plots for Gln89Ala system simulation showing the profiles of potential energy (A), root-mean-square deviation after the superposition of $\mathrm{C}_{\alpha}$ atoms to initial structure (B), gyration radius for all atoms (C) and root-mean square fluctuation of $\mathrm{C}_{\alpha}$ atoms from 80 to 100 ns (D).


Figure S4 - Diagnostic plots for Lys118Ala system simulation showing the profiles of potential energy (A), root-mean-square deviation after the superposition of $\mathrm{C}_{\alpha}$ atoms to initial structure (B), gyration radius for all atoms (C) and root-mean square fluctuation of $\mathrm{C}_{\alpha}$ atoms from 80 to 100 ns (D).


Figure S5 - Diagnostic plots for Asp247Ala system simulation showing the profiles of potential energy (A), root-mean-square deviation after the superposition of $\mathrm{C}_{\alpha}$ atoms to initial structure (B), gyration radius for all atoms (C) and root-mean square fluctuation of $\mathrm{C}_{\alpha}$ atoms from 80 to 100 ns (D).
-VIPR_ALG4_1084358903_577_2088 -VIPR_ALG4_AVZ47169_1_969_2480
-VIPR_ALG4_ASN64427_1_1_1512

- $\overline{\mathrm{VIPR}}$ _ALG4_AUI42289 $\overline{1} 1$ - 913 _2424

VIPR_ALG4_-ATW74960_1_1_1512
-VIPR_ALG4_ATWU_ASG45394_1_959_2470
-VIPR_ALG4_ASU55394_1_959_2470
-VIPR_ALG4_987031298_978_2489
पVIPR_ALG4_987031298_978_24
-VIPR_ALG4_1036637437_1_1512
-VIPR_ALG4_ASU55393_1_959_2470
[VIPR_ALḠ4_1046626550_969_2480
-VIPR_ALG4_1036637433_978_2489
-VIPR_ALG4_ASV60825
VIPR_ALG4_AUY62552_1_977_2488

- VIPR_ALG $\overline{4}$ _- 685052338 _977_2488
VIPR_ALG4_1036637431_977_2488
-VIPR_ALG4_ASV60828

| $\square$ VIPR_ALG4_ASV60828 |
| :---: |
| $\square$ VIPR_ALG4_ASV60827_1_1_1512 |
| VIPR_ALG4_ASV60826 |

VIPR_ALG4_592746960_871_2382
VIPR ALG4 592746966_871_2382
-VIPR_ALG4_ANŌ46307_1_918_2429
VIPR_ALG4_AMX81917-1_971_2482
-VIPR_ALG-ATW74961_1_1_1512
-VIPR_ALG4_1001229293_958_2469
[VIPR_ALG-G4_APH11536_1_961_2472
VIPR_ALG4_APH11534_1_962_2473
VIPR_ALḠ4_APH11539_1_962_2473 -VIPR_ALG4_APB03018_1_950_24 $\overline{4} 61$ -VIPR_ALG4_1115549286_871_2382 -VIPR ALG4 ASU55392 19072418 -VIPR ALG4 AWF93617-1960-2471 -VIPR ALG4-AOY08535-1-958-2469 -VIPR ALG4-985578256-971 2482 -VIPR_ALG4-985578256_971-2482 2468 -VIPR ALG4 AQS26816_1_957_2468 -VIPR_ALG4_AQS26816_1_896_2407 -VIPR_ALG4_AMK49165_1_964-2475
-VIPR_ALG4_AQS26791_1_896_2
-VIPR_ALG4_ARU07076_1_978_2489
-VIPR_ALG4_189092758_871_2382
-VIPR_ALG4_AMQ48982_1_978_2489
VIPR_ALG4_1087313262_871_2382
-VIPR_ALG4_APH11599_1_962_2473 -VIPR_ALG4_AOY08538_2_961_2472 -VIPR_ALG4 ASU55399_1 959 2470 -VIPR_ALG4_AQS26797_1_896-2407 -VIPR_ALG4_ARB07975-1-940-2451 -VIPR_ALG4_ARB07974_1_941_2452 -VIPR ALG4 ASU55411 - 959 - 470 -VIPR_ALG4-ASU55411_1_959_2470 -VIPR ALG4-984874584-871 2382 -VIPR_ALG4_984874584_871_2382
-VIPR_ALG4_973447409_917_2428
-VIPR_ALG4_ARU07182_1_871_2382
-VIPR_ALG4_1006593137_-970_2481
-VIPR_ALG4_APG56499_1_928_2439
-VIPR_ALG4_APB03022_1_950_2461
VIPR_ALG4_ARO85709_1_1_1512 [VIPR_ALG4_1020267436_950_246

VIPR_ALG4_AXE75586_1_1_1512
VIPR_ALG4_AXE75585_1 1 1512
-VIPR_ALG4_1029979472_1_1512
-VIPR ALG4 ASU55407 1-959 2470
-VIPR ALG4 ATG29301-1 962- 2473
-VIPR ALG4 ASL68979-1 962 - 2473
VIPR ALG4 100 $\overline{2} 39402 \overline{5}$ _946_245
VIPR ALG4 APG56457 18712382



 | |i| i|| IIT Ni













 Min


































 |||













Figure S6 - The phylogenetic tree generated with conservation analysis and the respective alignment matrix for complete sequences of ZIKV sE protein obtained from the Virus Pathogen Database and Analysis Resource (https://www.viprbrc.org/).


Figure S7 - Complexes formed by Z20 (green) and ZIKV-117 (cyano) on the ZIKV envelope protein showing the overlap of only the $\mathrm{E}(\mathrm{A})$ and Fab (B) regions, indicating the two antibodies are in different orientations in relation to protein $E$.


Figure S8 - Principal component analysis results showing the proportion of variance along the first 20 components (left) and the 2D distribution of the first two components (right).


Figure S9-Profiles of the distances (A) and angles (B) showing the dynamic of the cation- $\pi$ interaction between the residues Gln 89 A and Tyr 53 H . The distances were calculated between the hydrogen atoms 1 HE 2 (green) and 2HE2 (red) of the amino group of the side chain of Gln89A with the center of mass of the aromatic ring of the side chain of $\operatorname{Tyr} 53 \mathrm{H}$. The respective angles were calculated between the normal vector from the plane of the aromatic ring and the vector from the center of mass of aromatic ring to 1HE1 (green) and 2HE2 (red).

Table S1 - PDB codes of the structures whose sequences have 100\% identity and total coverage in relation to the consensus sequence of sE

| PDB id (chain) | Coverage | Identity (\%) | X-ray resolution |
| :---: | :---: | :---: | :---: |
| 5JHM (A e B) | 1.0 | 100 | $2.0 \AA$ |
| 5LBV (A) | 1.0 | 100 | $2.2 \AA$ |
| 5GZN (A) | 1.0 | 100 | $3.0 \AA$ |
| 5GZO (D) | 1.0 | 100 | $2.8 \AA$ |
| 5JHL (A) | 1.0 | 100 | $3.0 \AA$ |
| 5LBS | 1.0 | 100 | $2.2 \AA$ |

Table S2 - The contribution for van der Waals (vdW), electrostatic (Elec), polar solvation (Polar) and non-polar solvation (Nonpolar) of estimated free energies for the residue pairs ( $\Delta \mathrm{G}_{\text {pair }}$ ) selected with $\Delta \mathrm{G}_{\text {pair }}<-1 \mathrm{kcal}$.mol ${ }^{-1}$, the occurrence of hydrogen bonds showing between parenthesis the de number of individual donor-acceptor pairs, the normalized conservation scores for the EDII position and the residues variety at each position of the multiple sequence alignment. The energy terms are in kcal. $\mathrm{mol}^{-1}$

| EDII residue | $\mathbf{Z 2 0}$ residue | vdW | Elec | Polar | Nonpolar | $\Delta \mathrm{G}_{\text {pair }}$ | Occurrence (\%) | Normalized conservation score | Residue variety |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Ser64A | Asp102H | $0.2 \pm 0.8$ | $-5.8 \pm 1.3$ | $3.0 \pm 0.4$ | $-0.6 \pm 0.1$ | $-3.2 \pm 0.8$ | 87.0 (1) | 0.80 | Ser, Thr |
| Ser64A | Leu103H | $-0.9 \pm 0.2$ | $-0.2 \pm 0.1$ | $0.1 \pm 0.1$ | $-0.6 \pm 0.1$ | $-1.7 \pm 0.3$ | - |  |  |
| Ile65A | Leu103H | $-0.7 \pm 0.1$ | $0.2 \pm 0.1$ | $-0.2 \pm 0.1$ | $-0.4 \pm 0.1$ | $-1.1 \pm 0.2$ | - | 2.60 | Ile, Lys, Leu |
| Ile65A | Asn92L | $-0.6 \pm 0.3$ | $-1.1 \pm 1.0$ | $0.7 \pm 0.4$ | $-0.5 \pm 0.2$ | $-1.5 \pm 0.9$ | - |  |  |
| Ser66A | Leu103H | $-0.7 \pm 0.3$ | $-0.4 \pm 0.3$ | $0.4 \pm 0.2$ | $-0.8 \pm 0.1$ | $-1.5 \pm 0.3$ | - | -0.75 | Ser |
| Ser66A | His91L | $-0.7 \pm 0.5$ | $1.0 \pm 1.6$ | $-1.3 \pm 1.1$ | $-0.7 \pm 0.1$ | $-1.6 \pm 0.8$ | 0.2 (1) |  |  |
| Ser66A | Asn92L | $-0.9 \pm 0.2$ | $-0.2 \pm 0.3$ | $-0.1 \pm 0.2$ | $-0.5 \pm 0.1$ | $-1.7 \pm 0.4$ | - |  |  |
| Asp67A | His91L | $-0.6 \pm 0.1$ | $-18.7 \pm 1.1$ | $18.0 \pm 0.9$ | $-0.3 \pm 0.0$ | $-1.5 \pm 0.4$ | 75.4 (1) | -0.57 | Asp |
| Asp67A | Asn92L | $-0.7 \pm 0.2$ | $0.4 \pm 0.6$ | $-0.4 \pm 0.6$ | $-0.3 \pm 0.0$ | $-1.0 \pm 0.3$ | - |  |  |
| Asp67A | Tyr94L | $-1.3 \pm 0.2$ | $-2.5 \pm 0.4$ | $2.5 \pm 0.4$ | $-0.9 \pm 0.1$ | $-2.2 \pm 0.3$ | - |  |  |
| Asp67A | Arg96L | $-0.3 \pm 0.3$ | $-32.7 \pm 6.2$ | $30.0 \pm 2.5$ | $-0.3 \pm 0.2$ | $-3.3 \pm 4.0$ | - |  |  |
| Met68A | Asn92L | $-0.4 \pm 0.5$ | $-2.1 \pm 0.6$ | $0.3 \pm 0.3$ | $-0.7 \pm 0.1$ | $-2.9 \pm 0.6$ | 98.7 (1) | 2.53 | Ile, Met, Thr |
| Met68A | Ser93L | $-1.4 \pm 0.3$ | $-1.2 \pm 1.0$ | $0.2 \pm 0.5$ | $-1.0 \pm 0.1$ | $-3.3 \pm 0.5$ | - |  |  |
| Met68A | Tyr94L | $-0.7 \pm 0.3$ | $-1.2 \pm 0.5$ | $-0.1 \pm 0.2$ | $-0.4 \pm 0.1$ | $-2.5 \pm 0.3$ | - |  |  |
| Ala69A | Tyr94L | $-0.5 \pm 0.1$ | $-0.3 \pm 0.1$ | $0.2 \pm 0.0$ | $-0.6 \pm 0.1$ | $-1.2 \pm 0.2$ | - | 2.48 | Val, Ala, Thr |
| Lys84A | Asn59H | -0.1 $\pm 0.4$ | $-6.4 \pm 2.9$ | $4.3 \pm 1.7$ | $-0.6 \pm 0.2$ | $-2.9 \pm 1.3$ | 43.4 (1) | 1.44 | Arg, Lys |
| Gln89A | Trp34H | $-0.9 \pm 0.3$ | $-0.6 \pm 1.3$ | $0.2 \pm 0.8$ | $-0.8 \pm 0.1$ | $-2.1 \pm 0.8$ | 1.1 (1) | -0.57 | Gln |
| Gln89A | Tyr53H | $-1.1 \pm 0.2$ | $-0.7 \pm 0.5$ | $0.1 \pm 0.4$ | $-0.8 \pm 0.1$ | $-2.6 \pm 0.6$ | 0.1 (1) |  |  |

Table S2 (continued).

| EDII residue | Z20 residue | v.d.W | Elec | Polar | Non <br> Polar | $\Delta G_{\text {pair }}$ | Occurrence (\%) | Normalized conservation score | Residue variety |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Lys118A | Trp34H | $-1.6 \pm 0.2$ | $-2.4 \pm 0.9$ | $-0.6 \pm 0.7$ | $-1.4 \pm 0.1$ | $-5.9 \pm 1.0$ | 4.6 (1) | -0.51 | Lys |
| Lys118A | Glu51H | $0.8 \pm 0.7$ | $-53.5 \pm 1.8$ | $41.2 \pm 0.8$ | $-0.5 \pm 0.0$ | $-12.0 \pm 1.2$ | 98.2 (2) |  |  |
| Lys118A | Leu103H | $-1.0 \pm 0.3$ | $0.2 \pm 0.3$ | $-0.2 \pm 0.3$ | $-0.9 \pm 0.1$ | $-2.0 \pm 0.3$ | - |  |  |
| Ala 120A | Leu103H | $-0.6 \pm 0.5$ | $0.0 \pm 0.0$ | $0.0 \pm 0.0$ | $-0.7 \pm 0.1$ | $-1.3 \pm 0.5$ | - | 0.90 | Ala, Thr |
| Ser122A | Asp102H | $-0.6 \pm 0.2$ | $-0.3 \pm 1.1$ | $0.3 \pm 1.0$ | $-0.6 \pm 0.2$ | $-1.2 \pm 0.4$ | 0.1 (1) | 0.85 | Ser, Cys |
| Gly232A | His54H | $-0.3 \pm 0.3$ | $-3.0 \pm 1.8$ | $2.6 \pm 1.4$ | $-0.3 \pm 0.2$ | $-1.0 \pm 0.8$ | 3.3 (1) | -0.38 | Gly |
| Thr233A | Tyr53H | $-0.6 \pm 0.3$ | $-0.1 \pm 0.4$ | $0.0 \pm 0.1$ | $-0.5 \pm 0.2$ | $-1.2 \pm 0.6$ | 7.5 (1) | -0.70 | Thr |
| Lys246A | Arg30L | $-1.2 \pm 0.4$ | $28.6 \pm 3.0$ | $-27.9 \pm 2.8$ | $-1.0 \pm 0.3$ | $-1.6 \pm 0.5$ | - | 1.14 | Arg, Lys |
| Asp247A | Arg30L | $0.6 \pm 0.8$ | $-50.5 \pm 2.1$ | $39.9 \pm 0.9$ | $-0.8 \pm 0.1$ | $-10.8 \pm 1.3$ | 100.0 (3) | 1.05 | Asp, Glu |
| Arg252A | Gln27L | $-0.9 \pm 0.5$ | $0.4 \pm 2.9$ | $-0.5 \pm 2.3$ | $-0.9 \pm 0.4$ | $-1.8 \pm 1.2$ | 7.7 (3) | 1.10 | Arg, Thr |
| Arg252A | Gly28L | $0.0 \pm 0.4$ | $-4.4 \pm 3.6$ | $2.5 \pm 1.4$ | $-0.2 \pm 0.2$ | $-2.1 \pm 2.4$ | 37.0 (2) |  |  |
| Thr254A | Ile29L | $-0.6 \pm 0.2$ | $-0.2 \pm 0.2$ | $0.2 \pm 0.1$ | $-0.4 \pm 0.2$ | $-1.1 \pm 0.3$ | - | -0.70 | Thr |
| Thr254A | Arg30L | $-0.8 \pm 0.2$ | $2.2 \pm 0.7$ | $-1.7 \pm 0.5$ | $-0.7 \pm 0.1$ | $-1.1 \pm 0.4$ | 0.3 (1) |  |  |
| Val255A | Asn31L | $-0.1 \pm 0.3$ | $-1.6 \pm 1.2$ | $0.3 \pm 0.3$ | $-0.2 \pm 0.1$ | $-1.6 \pm 1.1$ | 56.1 (1) | 0.92 | Val, Ala |
| Val256A | Arg30L | $-0.8 \pm 0.2$ | $0.2 \pm 0.2$ | $-0.3 \pm 0.2$ | $-0.8 \pm 0.1$ | $-1.7 \pm 0.3$ | - |  |  |

