



OPEN

## In silico construction of a multiepitope Zika virus vaccine using immunoinformatics tools

Ana Clara Barbosa Antonelli<sup>1,6</sup>, Vinnycius Pereira Almeida<sup>1,6</sup>,  
Fernanda Oliveira Feitosa de Castro<sup>1,2</sup>, Jacycle Medeiros Silva<sup>3</sup>,  
Irmtraut Araci Hoffmann Pfrimer<sup>2</sup>, Edecio Cunha-Neto<sup>4,5</sup>, Andréa Queiroz Maranhão<sup>3,5</sup>,  
Marcelo Macedo Brígido<sup>3,5</sup>, Renato Oliveira Resende<sup>3</sup>, Anamélia Lorenzetti Bocca<sup>3</sup> &  
Simone Gonçalves Fonseca<sup>1,5</sup>✉

Zika virus (ZIKV) is an arbovirus from the Flaviviridae family and Flavivirus genus. Neurological events have been associated with ZIKV-infected individuals, such as Guillain-Barré syndrome, an autoimmune acute neuropathy that causes nerve demyelination and can induce paralysis. With the increase of ZIKV infection incidence in 2015, malformation and microcephaly cases in newborns have grown considerably, which suggested congenital transmission. Therefore, the development of an effective vaccine against ZIKV became an urgent need. Live attenuated vaccines present some theoretical risks for administration in pregnant women. Thus, we developed an *in silico* multiepitope vaccine against ZIKV. All structural and non-structural proteins were investigated using immunoinformatics tools designed for the prediction of CD4+ and CD8+ T cell epitopes. We selected 13 CD8+ and 12 CD4+ T cell epitopes considering parameters such as binding affinity to HLA class I and II molecules, promiscuity based on the number of different HLA alleles that bind to the epitopes, and immunogenicity. ZIKV Envelope protein domain III (EDIII) was added to the vaccine construct, creating a hybrid protein domain-multiepitope vaccine. Three high scoring continuous and two discontinuous B cell epitopes were found in EDIII. Aiming to increase the candidate vaccine antigenicity even further, we tested secondary and tertiary structures and physicochemical parameters of the vaccine conjugated to four different protein adjuvants: flagellin, 50S ribosomal protein L7/L12, heparin-binding hemagglutinin, or RS09 synthetic peptide. The addition of the flagellin adjuvant increased the vaccine's predicted antigenicity. *In silico* predictions revealed that the protein is a probable antigen, non-allergenic and predicted to be stable. The vaccine's average population coverage is estimated to be 87.86%, which indicates it can be administered worldwide. Peripheral Blood Mononuclear Cells (PBMC) of individuals with previous ZIKV infection were tested for cytokine production in response to the pool of CD4+ and CD8+ ZIKV peptide selected. CD4+ and CD8+ T cells showed significant production of IFN- $\gamma$  upon stimulation and IL-2 production was also detected by CD8+ T cells, which indicated the potential of our peptides to be recognized by specific T cells and induce immune response. In conclusion, we developed an *in silico* universal vaccine predicted to induce broad and high-coverage cellular and humoral immune responses against ZIKV, which can be a good candidate for posterior *in vivo* validation.

Zika virus (ZIKV) infection has been documented in 87 countries globally (by July of 2019)<sup>1</sup>. Brazil is the country with the highest number of ZIKV infection cases reported in the Americas, with the first registered autochthonous case in 2015<sup>2</sup>. The number of suspected and confirmed cases of autochthonous infections in Brazil reached more than 346 thousand people in the period of 2015 and March of 2017<sup>3</sup>. Even though the death rate is

<sup>1</sup>Department of Bioscience and Technology, Institute of Tropical Pathology and Public Health, Federal University of Goiás, Rua 235 s/n, sala 335, Setor Universitário, Goiânia, GO 74605-050, Brazil. <sup>2</sup>Department of Master in Environmental Sciences and Health, School of Medical, Pharmaceutical and Biomedical Sciences, Pontifical Catholic University of Goiás, Goiânia, Brazil. <sup>3</sup>Department of Cell Biology, University of Brasília, Brasília, Brazil. <sup>4</sup>Heart Institute (InCor), School of Medicine, University of São Paulo, São Paulo, Brazil. <sup>5</sup>Institute for Investigation in Immunology (iii) - National Institute of Science and Technology (INCT), São Paulo, Brazil. <sup>6</sup>These authors contributed equally: Ana Clara Barbosa Antonelli and Vinnycius Pereira Almeida. ✉email: [fonseca.simone@gmail.com](mailto:fonseca.simone@gmail.com)

low ( $<0.003\%$ )<sup>3</sup>, there were 14,558 suspected cases of microcephaly or congenital neurological malformations, from which 2,952 were confirmed<sup>4</sup>. Although the epidemic appeared to be waning after 2017, 31,576 cases were reported in the Americas in 2018, of which 3,589 were confirmed. In 2019, this number decreased to 12,079 reported cases and 3,482 confirmed cases<sup>1</sup>.

The ZIKV genome is composed by a positive-stranded RNA which codes a polyprotein that is subsequently cleaved into 3 structural proteins (Capsid C, prM, Envelope E) and 7 non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5). ZIKV is a mosquito-borne flavivirus<sup>5</sup> that has increased global attention in recent years due to the severe symptoms in certain populations. Other forms of transmission include blood transfusion, sexual, intrauterine, perinatal and laboratory manipulation. ZIKV-infected individuals are generally asymptomatic but can manifest rash, body itching, fever, headache, myalgia and arthralgia<sup>3</sup>. Abortion<sup>6</sup>, neurological malformation including microcephaly and other anatomical defects in fetus have also been associated with pregnant women being infected with ZIKV in 5–30% of cases<sup>7,8</sup>; in addition, in rare cases, Guillain-Barré syndrome (GBS), a neurodegenerative disease, has been reported after ZIKV infection<sup>9</sup>.

Protection against ZIKV is greatly associated with neutralizing antibodies against structural proteins, such as the Envelope E, particularly against the domain III<sup>10–12</sup>. As this humoral response is dependent on CD4 + T cells, such cell population is also required for complete protection against ZIKV infection, specially Th1 cells<sup>13–15</sup>. This finding was corroborated by Hassert and colleagues, who observed that in absence of CD4 + T cells, there is an increased disease severity in mice and higher viral titer in the nervous tissues<sup>16</sup>. Another important cell population in the control of ZIKV is CD8 + T cells, as demonstrated when IFN receptor-competent mice lacking CD8 + T cells had increased susceptibility to ZIKV pathologies<sup>17</sup> and transfer of CD8 + T cells from infected *Ifnar1*-/- mice to naïve mice lead to protection of the central nervous system after ZIKV challenge<sup>18</sup>. A screening of T cell epitopes revealed that the T cell response is directed mostly against structural proteins, such as prM and E, but also include epitopes from NS1, NS2A-B, NS3, NS4A-B and a high number of epitopes from NS5<sup>5</sup>. CD4 + T cells have been observed against all 10 ZIKV proteins, although most are directed towards the capsid and envelope proteins. CD8 + T-cell responses were mostly directed to NS3, NS4B and NS5 proteins<sup>19</sup>. These findings suggest that the correlates of protection in ZIKV infection are mediated by the complex interplay between specific B and T cells, providing insights into new strategies to guide protective immune responses against the infection.

Due to the risks related to the infection, including congenital malformations, prophylaxis against ZIKV infection is still an urgency, and the best answer is vaccinating people in risk areas, mostly susceptible populations. Pregnant women<sup>7</sup>, elderly<sup>20</sup> and immunocompromised individuals<sup>21</sup> deserve special attention. Several vaccine types against ZIKV are currently in phase I clinical trials, including DNA vaccines, mRNA-based, inactivated virus, peptide, and recombinant viral vectored vaccines<sup>22</sup>. However, hitherto only one DNA and one mRNA vaccine coding prM-E have succeeded to enter a phase 2 clinical trial, according to WHO vaccine pipeline tracker in April 2020 (<https://www.who.int/>). The most predominant form of antigens among them are the prM, E and NS1 proteins, but none of the listed vaccines are multiepitope vaccines or directed to other non-structural proteins<sup>23,24</sup>.

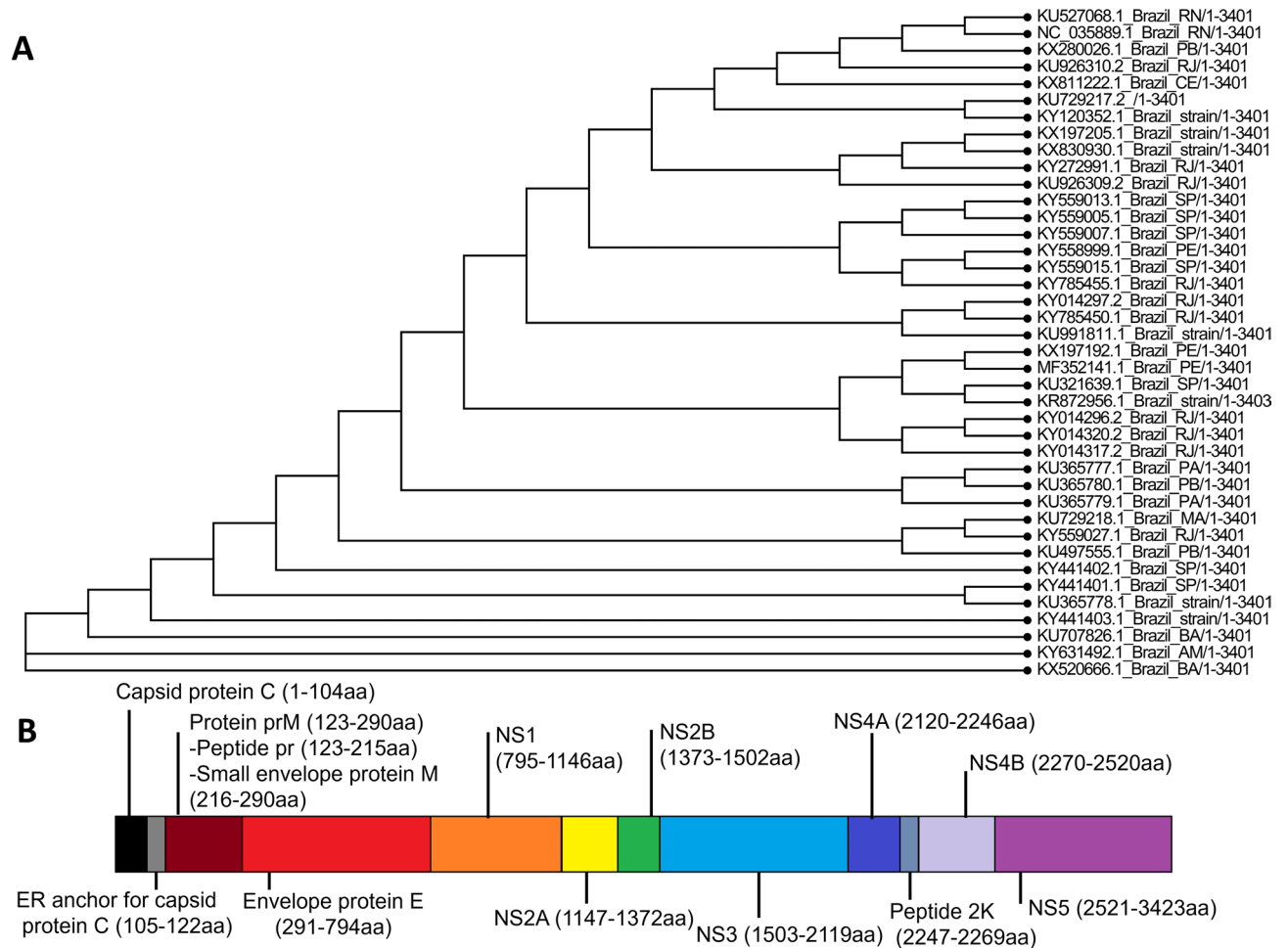
Multiepitope vaccines are composed of individual or overlapping epitopes flanked by linkers (small sequences of amino acids) and have attracted scientific interest because they are a safer alternative for pregnant women and other individuals at risk as they contain only portions of the pathogen proteins, and therefore cannot cause infection. Furthermore, they are able to induce long-lasting humoral and cellular immunity, have the potential to include a repertoire of epitopes with potential to bind a high variety of Human Leukocyte Antigens (HLA) types, and therefore a high population coverage<sup>25</sup>. Multiepitope vaccines containing both B and T cells epitopes against Hepatitis C virus (HCV) have been tested in animal models and demonstrated increase in antibodies and specific T cell response<sup>26</sup>, and an HIV vaccine containing multiple CD8 + T cell epitopes has entered clinical trial phase I (NCT00076037)<sup>27</sup>. Furthermore, there are several multiepitope vaccines in clinical trials phase I and II for neoplasia, including melanoma (NCT00580060, NCT00153569, NCT00003362), breast cancer (NCT03793829, NCT03012100, NCT04144023), glioblastoma<sup>28</sup>, prostate (NCT02362464, NCT02362451) and ovarian cancer (NCT02764333) (<https://clinicaltrials.gov/>). Altogether, the safety profile and the success of multiepitope vaccines in entering clinical trials are encouraging features for a vaccine against ZIKV.

Aiming to develop such vaccine, we performed extensive literature and bioinformatics analysis to select promising ZIKV protective epitopes derived from all structural and non-structural proteins, seeking promiscuous sequences that cover the highest number of type I and type II HLAs. We assembled a novel *in silico* chimeric ZIKV vaccine (ZIKVac), composed of three main portions. At the N' terminal, there is the ZIKV Envelope protein domain III (EDIII), followed by CD4 + T cell epitopes and finally by overlapping CD8 + T cell epitopes. The vaccine was evaluated with different protein adjuvants at the C' terminal. Analysis of the biochemical and stability parameters were performed, as well as safety and immunogenicity indicators.

## Results

**ZIKV polyprotein consensus sequence generation.** Aiming the selection of conserved epitopes, a consensus sequence was generated from 40 complete Brazilian ZIKV sequences available on NCBI (GenBank access number on Supplementary 1). A Phylogenetic tree of all 40 complete sequences was generated, showing high conservancy between them (Fig. 1A). This sequence was later separated into single proteins, which were used as input in FASTA format for the selection of potential type I and II immunogenic epitopes. Figure 1B portrays the obtained ZIKV polyprotein consensus sequence and shows the start and end amino acids of each structural and non-structural protein.

**Selection of CD4 + and CD8 + T cell epitopes.** The exploration of multiple epitope prediction servers to identify good CD4 + and CD8 + T cell epitopes may decrease the rate of false positive targets for vaccine construction. Here, at least 3 different platforms based on non-related prediction algorithms were compared. Pro-



**Figure 1.** Phylogenetic tree of the 40 complete ZIKV strains used for obtaining the consensus sequence (A). (B) Polyprotein of the ZIKV from the Brazilian consensus sequence. The ZIKV genome codes for seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) and four structural proteins (Capsid protein C, Protein prM, ER anchor for capsid protein C and Envelope protein E). Protein prM is further digested into Peptide pr and Small envelope protein M. ER anchor for capsid protein C is removed by serine protease NS3. The consensus sequence was obtained from 40 complete Brazilian sequences deposited on NCBI.

pred1 and Propred were used to identify the most promiscuous CD8+ and CD4+ T cell epitopes, respectively. We also used the Binding affinity and Immunogenicity tools available on IEDB, and for CD4 epitopes, we also evaluated IFN- $\gamma$  induction. We first prioritized epitopes that scored above the desired threshold in all servers, however, a considerable part of the selected epitopes was considered hydrophobic and therefore not suitable for synthesis and purification. We then selected only non-hydrophobic peptides, which considerably restricted our options. Hence, it was not possible to select only peptides for all ZIKV proteins with high scores in all servers, so some parameters were prioritized, such as promiscuity. The selected CD8 and CD4 peptides and their HLA coverage, binding affinity score and immunogenicity score are summarized on Tables 1 and 2, respectively.

ZIKV proteins are poor in CD8 epitopes when compared with CD4 epitopes, as predicted by both Propred1 and the IEDB MHC-I binding prediction server. CD8 epitopes are also inferior when it comes to HLA coverage. Because of that, a threshold of 3% was used for CD8 epitopes on Propred1, while CD4 peptides were selected with a 1% threshold. Moreover, to increase HLA coverage of all ZIKV proteins, more than one peptide was selected for some proteins. NS4A CD8 epitopes GKMFGMVTLG (NS4A79-89) and GAAFGVMEALGT (NS4A1-12) alone, for example, cover 22.5% and 30% of the available HLA, respectively, but together they cover 52.5%, an average coverage for a CD8 epitope. CD4 epitopes, on the other hand, presented greater HLA coverage and binding affinity in general. Also, 9 out of 12 CD4 peptides were shown to induce IFN- $\gamma$  production (Table 2).

A total of 13 CD8+ and 12 CD4+ T cell epitopes that covered all structural and non-structural ZIKV proteins were selected for the vaccine construction.

Some of the CD4+ and CD8+ T cell epitopes we selected presented negative scores for immunogenicity (E\_415-425) and IFN induction (prM\_71-85, E\_136-150, E\_296-310). However, these epitopes are important for the population coverage of our vaccine, so we performed additional analysis to make sure these epitopes would not bring negative outcomes for the whole vaccine immunogenicity. Therefore, we evaluated the capacity of the mentioned epitopes to induce IL-10 and IL-4, to activate APCs, and to induce proinflammatory response. None of these 4 epitopes were considered inducers of the immunoregulatory cytokines IL-10 and IL-4 and all of them

Protein	Epitope <sup>0</sup>	Size (aa)	Position	N° of alleles <sup>1</sup>	% of total alleles <sup>2</sup>	Immunogenicity score (IEDB) <sup>3</sup>	N° of HLAs w/ Percentile Rank $\leq 20$ <sup>4</sup>	N° of HLAs with SMM IC50 $\leq 500$ nM <sup>5</sup>	N° of HLAs with ANN IC50 $\leq 500$ nM <sup>6</sup>
Capsid Protein	<u>V</u> SPFGGLKRLP (C_24-34)	11	24–34	17	42.5	0.0036	12	1	1
prM Protein	<u>I</u> FRNPGFA-LAAAAIA (prM_129-143)	15	129–143	16	34.04	0.3976	49	4	8
Envelope Protein E	<u>S</u> QEGAVHTALA (E_260-270)	11	260–270	10	25	0.23948	15	3	3
Envelope Protein E	<u>V</u> RGAKRMAVLG (E_415-425)	11	415–425	10	25	-0.21504	15	2	1
Envelope Protein E	<u>M</u> QTLTPVGRLLI (E_349-359)	11	349–359	9	22.5	0.13774	26	3	2
NS1	<u>W</u> RSVEGEL- <u>N</u> AILEENGVQL (NS1_68-86)	19	68–86	11	23.40	0.5434	34	3	3
NS2A	<u>A</u> LIAAFKVR (NS2A_78-86)	9	78–86	5	10.64	0.07299	8	1	1
NS2B	<u>S</u> EVLTAVGLICAL (NS2B_5-17)	13	5–17	22	46.81	0.24318	40	8	7
NS3	<u>V</u> REAIKTRLRT-VIL (NS3_209-222)	14	209–222	14	29.79	0.26841	34	6	6
NS4A	<u>G</u> KMGFGM-VTLG (NS4A_79-89)	11	79–89	9	22.5	0.01824	24	4	4
NS4A	<u>G</u> AAFG-VMEALGT (NS4A_1-12)	12	1–12	12	30	0.17747	29	4	3
NS4B	<u>T</u> PLTLIVAI (NS4B_105-113)	9	105–113	13	27.66	0.20764	10	2	1
NS5	<u>Y</u> MWLGAR-FLEFEAL (NS5_477-490)	14	477–490	17	36.17	0.6005	54	6	6

**Table 1.** Zika virus CD8 peptide selection based on multiparameters using immunoinformatics tools. 0—Underlined are the first aminoacids of a 9-mer peptide predicted as having percentile rank  $\leq 20$  as calculated by IEDB (<http://tools.iedb.org/mhci/>). 1—Total number of HLA-I covered by the peptide as predicted by Propred1. All 47 alleles were selected, and the percentage of the highest scoring peptides, the Immunoproteasome and the Proteasome filter threshold was set up as 3%. 2—Percentage of HLA-I molecules identified as binders in relation to the total number of HLA-I available. 3—Immunogenicity score obtained with <http://tools.iedb.org/immunogenicity/>. Positive values are considered immunogenic. 4—Number of HLA that presented percentile rank lower than 20 as predicted by IEDB (<http://tools.iedb.org/mhci/>). 5—Number of HLAs with SMM IC50  $\leq 500$  nM as predicted by IEDB (<http://tools.iedb.org/mhci/>). 6—Number of HLAs with ANN IC50  $\leq 500$  nM as predicted by IEDB (<http://tools.iedb.org/mhci/>).

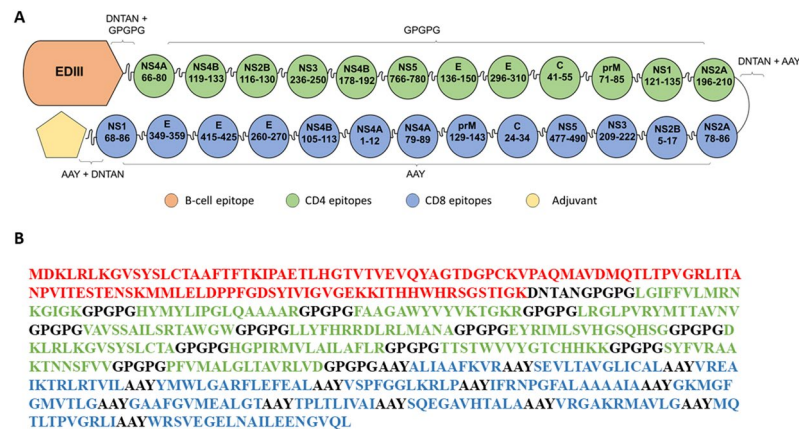
were considered capable of activating APCs. Moreover, all the four epitopes were considered proinflammatory. These results are summarized on Supplementary 3.

**Vaccine construction.** We rearranged CD4+ and CD8+ T cell epitopes, separated them by proper linkers in order to elect the best vaccine structure based on allergenicity, antigenicity and autoimmunity analyses, secondary and tertiary structure quality and discontinuous epitope scores. EDIII was added to the structure in order to induce B cell response. The DNTAN linker is a loop in protein (LiP) and is appropriate to separate BCR epitopes as there is a need of flexibility between epitopes so they can be recognized directly by B-cells<sup>29</sup>; therefore, we used it to separate EDIII from the rest of the protein. GPGPG and AAY motifs were used to flank CD4 and CD8 epitopes, respectively. We evaluated seven designs (Supplementary 4) regarding their quality of tertiary structure, physical-chemical parameters and antigenicity. We opted to continue the analysis with the structure 6 because it presented the best antigenicity while being predicted to be stable and to have a high expected half-life (Supplementary 5). The EDIII tertiary structure of the final vaccine construct had a high overlap (RMSD 0.845 angstrom) with the EDIII from the crystal structure of ZIKV Envelope protein (PBD: 5JHM), suggesting that the addition of the CD4+ and CD8+ T cell epitopes to the protein structure did not impact the EDIII folding (Supplementary 6). We also added different adjuvants to the vaccine construct to perform our analysis. The final construct (ZIKVac) is shown in Fig. 2 and the ZIKVac sequence associated with adjuvants can be seen in Supplementary 7.

**Evaluation of vaccine autoimmunity induction.** Only Flagellin and HBH had sequences with more than six identical amino acids. Flagellin has eight identical amino acids to the human Myosin VI isoform 3, and

Protein	Epitope	Size (aa)	Position	N° of alleles <sup>1</sup>	% of total alleles <sup>2</sup>	Immunogenicity (Percentile Rank) <sup>3</sup>	N° of HLAs w/Percentile Rank < 20 <sup>4</sup>	N° of HLAs w/ SMM align IC50 < 500 nM <sup>5</sup>	N° of HLAs w/NN align IC50 < 500nM <sup>6</sup>	IFNepitope Score (SVM) <sup>7</sup>
Capsid Protein	HGPIRMVLAAILAFLR (C_41-55)	15	41–55	37	72.54	24.3268	21	17	14	0.4521
prM Protein	TTSTWVVYGTCH-HKK (prM_71-85)	15	71–85	31	60.8	65.45964	4	0	2	-0.1269
Envelope Protein E	EYRIMLSVHGSQHS (E_136-150)	15	136–150	33	64.7	40.45516	13	9	12	-0.1841
Envelope Protein E	DKLRLKGVSYSLCTA (E_296-310)	15	296–310	34	66.6	46.33096	12	2	12	-0.3609
NS1	SYFVRAAKTNSNFVV (NS1_121-135)	15	121–135	30	58.8	31.18672	11	12	12	0.3363
NS2A	PFVVALGLTAVRLVD (NS2A_196-210)	15	196–210	29	56.8	42.90484	17	11	14	0.2505
NS2B	FAAGAWYVYVKTGKR (NS2B_116-130)	15	116–130	11	27.50	42.51056	14	5	9	0.5012
NS3	LRGLPVRVMTTAVNV (NS3_236-250)	15	236–250	35	68.6	43.8254	12	8	13	0.4513
NS4A	LGIFFVLMRNKGIGK (NS4A_66-80)	15	66–80	40	78.4	25.47468	21	10	14	0.0979
NS4B	HYMYLIPGLQAAAAAR (NS4B_119-133)	15	119–133	18	35.29	42.61168	19	10	14	0.0579
NS4B	VAVSSAILSRTAWGW (NS4B_178-192)	15	178–192	38	74.5	49.2798	13	6	6	0.3106
NS5	LLYFHRRLRLMANA (NS5_766-780)	15	766–780	36	70.6	38.39304	20	8	12	0.2457

**Table 2.** Zika virus CD4 peptide selection based on multiparameters using immunoinformatics tools. 1—Total number of HLA-II covered by the peptide as predicted by Propred MHC-II Binding Peptide Prediction Server. All 51 alleles were selected. The threshold was set up as 1%. 2—Percentage of HLA-II molecules identified as binders in relation to the total number of HLA-II available. 3—Immunogenicity percentile rank obtained with <http://tools.iedb.org/CD4episcore/>. Lower numbers are considered immunogenic (lower than 66). 4—Number of HLA that presented percentile rank lower than 20 (high affinity). 5—Number of HLA that presented SMM IC50 lower than 500 nM (high affinity). 6—Number of HLA that presented NN IC50 lower than 500 nM (high affinity). 7—IFN-γ induction score. Positive values indicate IFN-γ induction.



**Figure 2.** ZIKVac structure showing the position of CD4 (green) and CD8 epitopes (blue) envelope protein E Domain III (EDIII) and adjuvant (yellow pentagon) (A). The protein sequence of and ZIKVac sequence without conjugated adjuvants showing the EDIII (red), CD4 epitopes (green), CD8 epitopes (blue) and linkers (black) (B). DNTAN linkers were used to separate EDIII and adjuvant from the epitopes, GPGPG linkers were used between CD4 epitopes and AAY linkers between CD8 epitopes.

HBH presents 8 identical amino acids to Histone H1.5. ZIKVac presented sequences with a maximum of only five identical amino acids compared to human proteins (Supplementary 8).

**Allergenicity, antigenicity and physicochemical parameters prediction of ZIKVac.** ZIKVac alone and associated with each of the evaluated adjuvants (Flagellin, 50S Rib, HBH and RS09) were considered non-allergenic by AllergenFP server (Table 3). Antigenicity of all structures was evaluated using Vaxijen server,



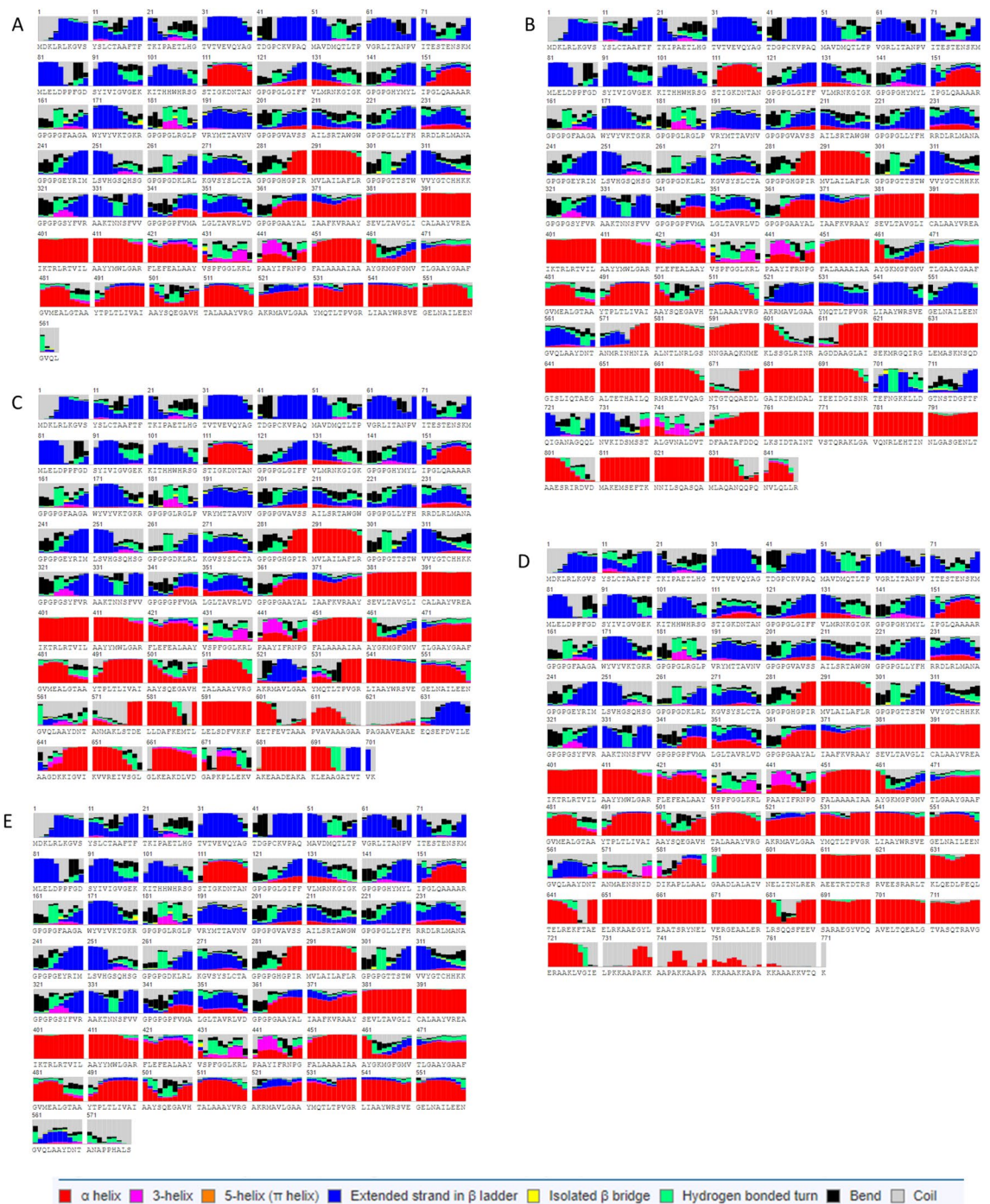
Properties	Server	ZIKVac	ZIKVac + flagellin	ZIKVac + HBH	ZIKVac + RS09	ZIKVac + 50S Rib
Length (aa)	ProtParam ExPASy Server	564	847	771	579	702
Probable non-allergen	AllergenFP Server	+	+	+	+	+
Probable antigen Virus, Thr: 0.4	Vaxijen Server	+(0.4749)	+(0.5219)	+(0.4427)	+(0.4751)	+(0.4580)
Stable (Instability index)	ProtParam ExPASy Server	+(25.75)	+(26.81)	+(31.20)	+(26.27)	+(24.72)
Secondary structure	RaptorX Structure Prediction Server	32%H, 26%E, 40%C	39%H; 22%E; 37%C	41%H; 18%E; 39%C	31%H; 26%E; 41%C	33%H; 24%E; 42%C
Solvent accessibility	RaptorX Structure Prediction Server	41%E, 24%M, 34%B	43%E; 25%M; 30%B	45%E; 22%M; 32%B	42% E; 23%M; 33%B	43%E; 24%M; 32%B
Theoretical pI	ProtParam ExPASy Server	9.82	9.4	9.65	9.77	9.29
Molecular weight	ProtParam ExPASy Server	58.78 kDa	89.06 kDa	81.12 kDa	60.28 kDa	73.02 kDa
Negatively charged residues	ProtParam ExPASy Server	26	60	59	27	52
Positively charged residues	ProtParam ExPASy Server	51	77	87	51	67
Extinction coefficients at 280 nm (if all pairs of Cys residues form cystines)	ProtParam ExPASy Server	78,980 M <sup>-1</sup> cm <sup>-1</sup>	80,470 M <sup>-1</sup> cm <sup>-1</sup>	84,940 M <sup>-1</sup> cm <sup>-1</sup>	80,470 M <sup>-1</sup> cm <sup>-1</sup>	80,470 M <sup>-1</sup> cm <sup>-1</sup>
Extinction coefficients at 280 nm (if all pairs of Cys residues are reduced)	ProtParam ExPASy Server	78,730 M <sup>-1</sup> cm <sup>-1</sup>	80,220 M <sup>-1</sup> cm <sup>-1</sup>	84,690 M <sup>-1</sup> cm <sup>-1</sup>	80,220 M <sup>-1</sup> cm <sup>-1</sup>	80,220 M <sup>-1</sup> cm <sup>-1</sup>
Estimated half life	ProtParam ExPASy Server	30 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (Escherichia coli, in vivo)	30 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (Escherichia coli, in vivo)	30 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (Escherichia coli, in vivo)	30 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (Escherichia coli, in vivo)	30 h (mammalian reticulocytes, in vitro) > 20 h (yeast, in vivo) > 10 h (Escherichia coli, in vivo)
Aliphatic index	ProtParam ExPASy Server	89.01	88.02	87.35	88.24	91.15
GRAVY (Grand average of hydropathicity)	ProtParam ExPASy Server	0.233	0.013	0.008	0.215	0.210

**Table 3.** Allergenicity, antigenicity and physicochemical parameters of the ZIKVac and with association with the proteinaceous adjuvants.

and antigenicity of ZIKVac was predicted to be 0.4749 (Table 3). Interestingly, ZIKVac associated with adjuvants 50S Rib and HBH presented lower antigenicity scores than ZIKVac alone (0.4580 and 0.4427 respectively), while RS09 did not interfere with the vaccine's antigenicity levels (0.4751). In contrast, flagellin increased ZIKVac's antigenicity considerably (0.5219). Isoelectric point (pI) ranged between 9.4 and 9.85, which indicates that all constructs are basic in nature. The instability index (II), aliphatic index (AI) and GRAVY were very similar between all structures, meaning that all of them are stable (II < 40), probably thermostable (high AI), and slightly hydrophobic (GRAVY > 0) (Table 3).

**Secondary structure.** The secondary structure as predicted by RaptorX reveals that ZIKVac is composed of 32% alpha helices, 26% beta-strands and 40% of coils (Fig. 3). In respect of solvent accessibility, 41% of the protein is exposed, 24% medium and 34% buried. In combination with the adjuvants, the resulting secondary structure varied as shown in Table 5. ZIKVac + flagellin and ZIKVac + HBH contained the highest percentage of alpha helix and the lowest percentage of beta-sheet conformation. The amount of coil formation was about the same in all structures, ranging from 37 to 42%. The addition of the adjuvant did not change the rate of solvent accessibility within the protein.

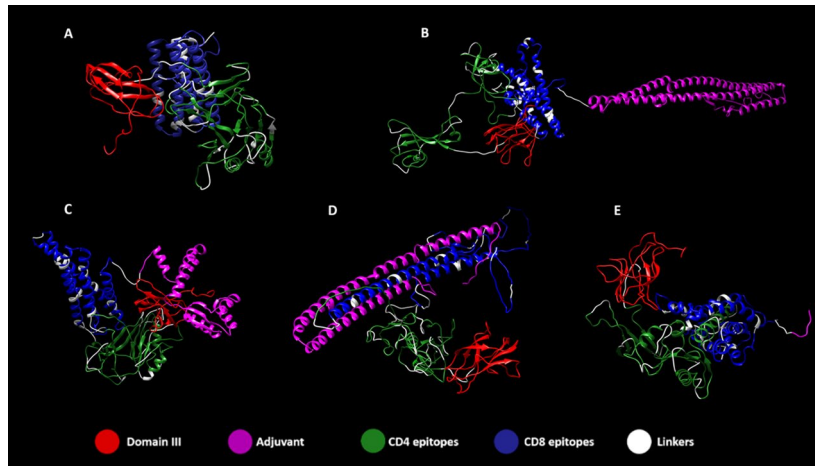
**Tertiary structure refinement and validation.** Tertiary structure of proteins was obtained with RaptorX and refined with 3D Refine, followed by validation with RAMPAGE. Figure 4 displays 5 vaccine structures of ZIKVac and ZIKVac associated with the four different adjuvants. Interestingly, in most cases, domain III (Fig. 4A,B,D,E) was separated from the rest of the protein structure, which facilitates antibody recognition and possible attachment with innate cellular receptors, facilitating vaccine intake by the cells. The CD4 and CD8 epitope-rich regions of the protein acquired different conformations according to the adjuvant added. We placed EDIII at the N-terminus to minimize epitope masking by neighbouring domains allowing the greatest number of potential BCRs epitopes to be available. Further, positioning EDIII centrally in the construct may reduce potential effects from neighbouring domains on EDIII folding. The same concept applies for the positioning of the adjuvants at the C-terminus as it can increase the potential TLR binding.



**Figure 3.** Secondary structure of the vaccine protein alone (A) and in association with adjuvants (B–E). Helices, coils and beta-sheets are specified in the legend, to note: alpha helix in red, 3-helix in magenta, 5-helix (pi helix) in orange, extended strand in beta ladder in blue, isolated beta bridge in yellow, hydrogen bonded turn in green, bend in black and coil in gray. ZIKVac + Flagellin (B), 50S ribosomal (C), HBH (D), RS09 (E).

Validation of refined tertiary structures showed that all constructs had between 91.0 and 92.7% of residues in favored regions. ZIKVac, ZIKVac + Flagellin and ZIKVac + 50S Rib presented 2.5%, 2.8% and 2.9% of residues in outlier regions, respectively, while HBH and RS09 increased the percentage of residues in outlier regions to 3.6% and 3.3%, respectively (Fig. 5).

**Continuous and discontinuous B-cell epitopes visualization.** Prediction of conformational B cell epitopes from the refined model of ZIKVac were obtained with Ellipro and revealed ten discontinuous epitopes (Fig. 6) and fourteen continuous B cell epitopes (Table 4). The three best continuous and the best and third best discontinuous epitope scores were located at EDIII, corroborating with previous findings that this domain is a good antibody inducer.



**Figure 4.** Tertiary structure of ZIKVac (A) with highlighted epitopes flanked by linkers and the adjuvants at the C-terminal position, which are the flagellin (B), 50S Ribosomal (C), Heparin-binding hemagglutinin (HBH) (D) and RS09 (E). In red is the ZIKV Envelope Domain III, in magenta the adjuvants, in green CD4 epitopes, in blue CD8 epitopes and in white the linkers.

**Population coverage.** In order to confirm whether our vaccine could be effective in different regions and populations, we used the Population Coverage tool available on IEDB online. Only ZIKVac CD4 and CD8 epitopes were considered, not the whole protein construct. It was observed high coverage in several world regions, as indicated on Table 5. Our vaccine's worldwide population coverage according to the IEDB server was 96.81%, and the highest coverage was in South Asia (98.71%), while the lowest was in South Africa (18.36%). In contrast to South Africa, Central, North, East and West Africa showed 95.44%, 77.17%, 96.18% and 97.95% coverage, respectively. In Central and South America population coverage by our peptides was 92.54% and 94.9%, respectively.

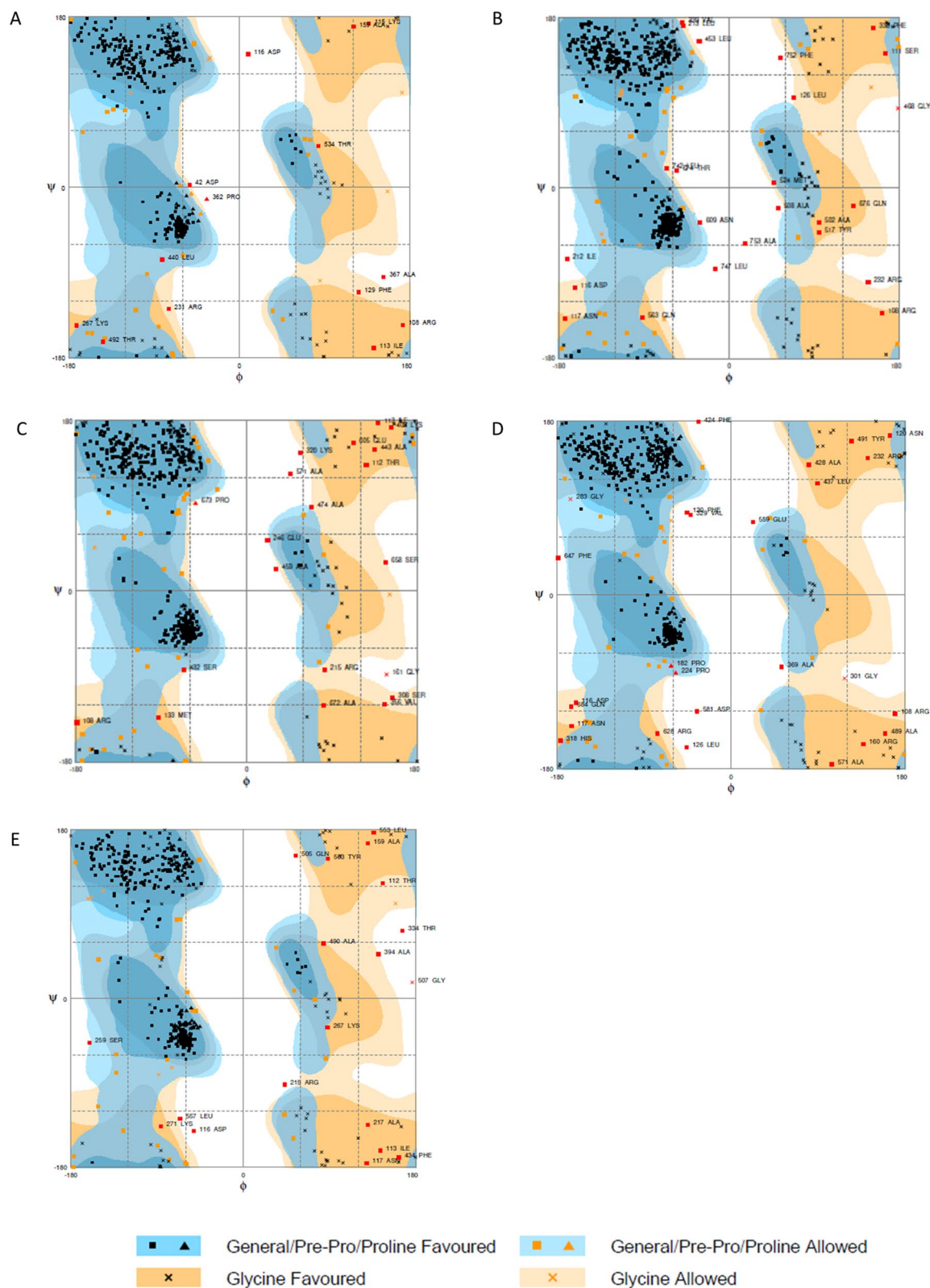
**Detection of cytokine-producing CD4 + and CD8 + T cells in response to ZIKV peptides in individuals with previous ZIKV infection.** In order to confirm that the predicted peptides could be recognized by T cells from individuals with previous infection with ZIKV and able to induce immune response, we stimulate PBMC of individuals with previous exposure to ZIKV with the pool of 13 CD8 and 12 CD4 peptides selected for the ZIKV vaccine. Cytokine production by CD4 + and CD8 + T cells was assessed after stimulation using flow cytometry (Fig. 7). Gate strategy is represented in Supplementary 9. Both CD4 + and CD8 + T cells showed significant production of IFN- $\gamma$  (7A, 7D). IL-2 production was also observed by CD8 + T cells (7E), showing a polyfunctionality of CD8 + T cells responses to these epitopes. Some individuals also showed higher TNF- $\alpha$  levels after stimulation, however there was no statistical difference between stimulated and non-stimulated cells for this cytokine (7C, 7F).

## Discussion

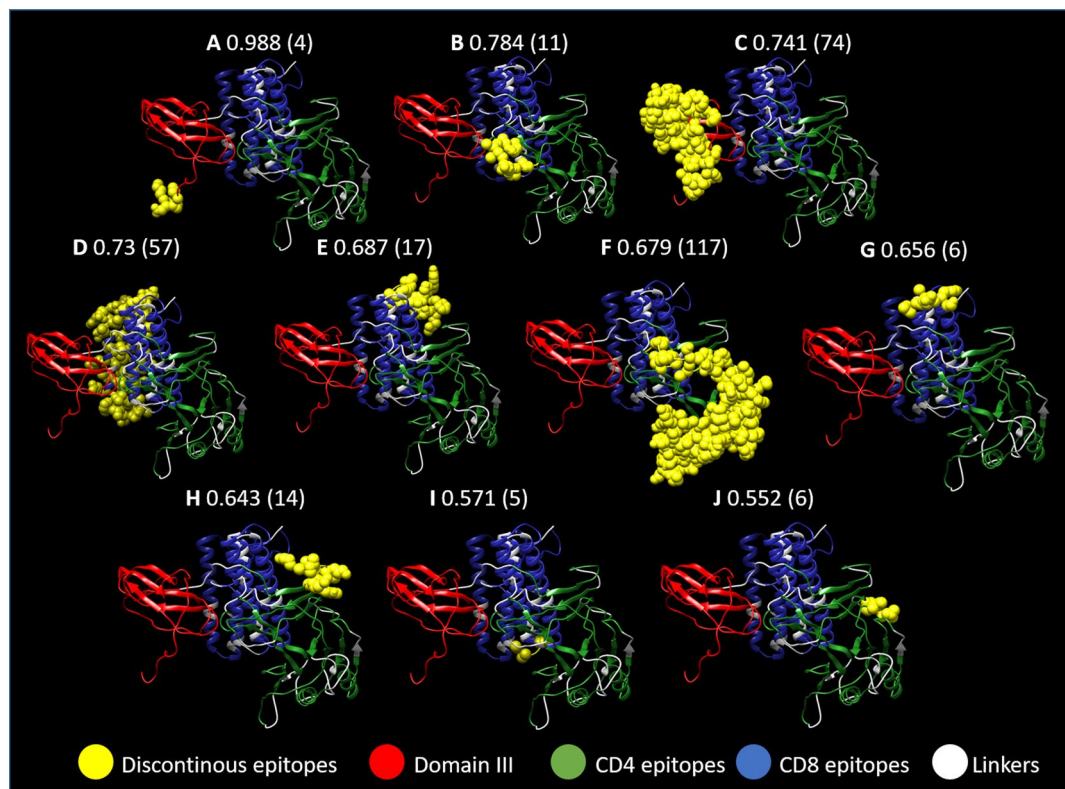
ZIKV infection was considered a public health concern in tropical countries, especially in South America due to a rapid growth in the number of microcephaly cases in Brazil caused by the flavivirus in 2015. Although the number of reported cases has waned in the last two years, it could increase anytime with the spread of the mosquito vector during rainy seasons<sup>30</sup>. Considering the potential risks of such infection, it is clear the need of a prophylactic vaccine against it. As the infection is a danger to the fetus, pregnant women in their first trimester would be the main target of such vaccination, as this is the period in which the virus is known to cause more severe neurological damage to the fetus<sup>31,32</sup>. Considering this population, we designed a chimeric multiepitope vaccine that should stimulate the production of specific antibodies, CD4 + and CD8 + T cells to control and prevent viral spread without risk of infection. For selecting the epitopes, we used a consensus sequence derived from all 40 complete Brazilian ZIKV strains available on NCBI by March of 2018. Interestingly, a later analysis showed that this consensus sequence would be the same if all ZIKV genomes obtained in American countries by March of 2018 were considered (about 316 complete and incomplete genomes).

For an effective humoral response, it is fundamental that the vaccine is capable of inducing neutralizing antibodies that lead to inhibition of viral replication, protect against the infection and disease. Flaviviruses structural proteins, especially the Envelope protein, are the main targets for antibody production<sup>33,34</sup>. The ZIKV EDIII has been used in experimental subunit vaccines, and it was shown to induce neutralizing antibodies against different strains of ZIKV, as well as significant reduction of viral replication when used in combination with alum and Monophosphoryl lipid A (MPL)<sup>35</sup>. Conversely, ZIKV EDIII failed to induce satisfactory inhibition of viral replication in mice when given in the form of DNA, protein or when expressed by a chimpanzee adenovirus<sup>36</sup>, indicating that the choice of adjuvant, for instance, can impact the efficacy of the vaccine. Interestingly, immunization with a specific fragment of EDIII (E298-409) was sufficient to protect BALB/c pups against lethal challenge with ZIKV strains R103451 and FLR<sup>37</sup>.





**Figure 5.** Ramachandran plots of the ZIKVac alone (A) and linked with the adjuvants, which are the flagellin (B), 50S Ribosomal protein (C), Heparin-binding hemagglutinin (HBH) (D) and RS09 (E). Black squares and triangles in blue areas indicate General/Pre-Pro/Proline Favoured, orange squares and triangles in lighter blue areas indicate General/Pre-Pro/Proline Allowed, black Xs in orange regions indicate Glycine Favoured and Orange Xs in lighter orange regions indicate Glycine allowed.



**Figure 6.** Discontinuous epitopes present in the vaccine construct. Two epitopes are located at the Envelope Domain III portion of the protein (A and C) with some of the highest scores (0.988 and 0.741) and with 4 and 74 amino acids, respectively. The other discontinuous epitopes are distributed among the CD4 and CD8 epitopes (B, D–J). In yellow are the predicted discontinuous epitopes, in red the domain III from the ZIKV Envelope protein, in green CD4 epitopes, in blue CD8 epitopes and in white the linkers.

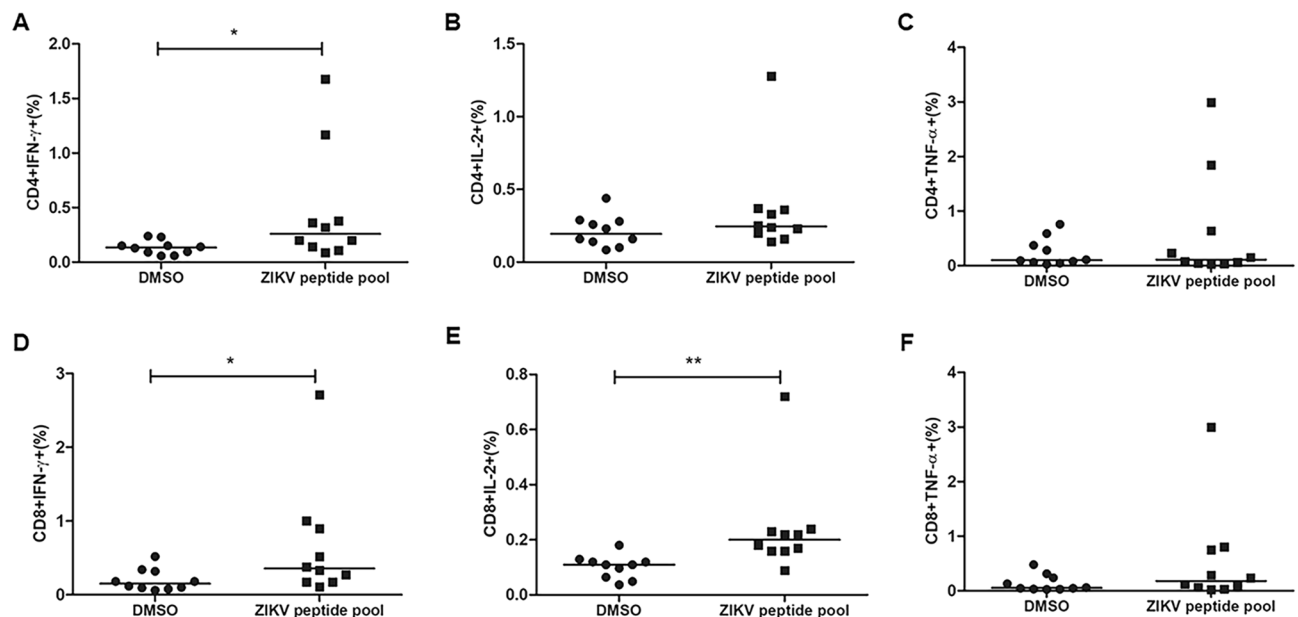
ZIKVac domain	Start–end	Peptide	Number of residues	Score
Domain III	83–94	ELDPPFGDSYIV	12	0.836
Domain III	1–13	MDKLRLKGVSYSL	13	0.777
Domain III	45–69	CKVPAQMAVDMQTLTPVGRLLITANP	25	0.774
CD4 epitope domain	154–164	LQAAAARGPGP	11	0.775
Domain III	25–32	AETLHGTV	8	0.739
CD4 epitope domain	200–245	VGPGPGVAVSSAILLSRTAWGWGPGPGLLYFHRRDLRLMANAG-PGPG	46	0.734
CD4 epitope domain	281–325	GPGPGHGPIRMVLAFLRGPGGTTSTWVVYGTCHHK-KGPGPG	45	0.716
CD8 epitope domain	517–564	YVRGAKRMAVLGAAYMQTLTPVGRLLIAAYWRSVEGELNAI-LEENGVQL	48	0.715
Domain III	96–113	GVGEKKITHHWHRSGSTI	18	0.7
CD8 epitope domain	434–450	FGGLKRLPAAYIFRNPG	17	0.687
CD4 epitope domain	134–145	RNKGIGKGP	12	0.658
CD4 epitope domain	253–265	VHGSQHS GGPGPG	13	0.635
CD4 epitope domain	183–187	GPGLR	5	0.567
CD8 epitope domain	471–476	TLGAAY	6	0.564

**Table 4.** Predicted Linear B cell Epitopes from ZIKVac.

With that in mind, we chose ZIKV EDIII to cover our vaccine's B cell response. EDIII from Dengue virus and other flaviviruses have been previously reported to induce neutralizing antibodies<sup>38–40</sup>. Similarly, ZIKV EDIII specific antibodies have also shown neutralizing activity, while domains I and II have presented either modest or none<sup>11,41,42</sup>. EDIII was therefore a promising choice for a vaccine construct. The inclusion of the ZIKV EDIII to the multi-epitope vaccine structure is a novelty of our work.

Population/Area	Class combined		
	Coverage <sup>a</sup>	Average_hit <sup>b</sup>	pc90 <sup>c</sup>
Central Africa	95.44%	50.89	29.93
Central America	92.54%	42.49	26.86
East Africa	96.18%	57.81	32.85
East Asia	89.15%	43.81	23.02
Europe	98.07%	73.1	48.88
North Africa	77.17%	31.37	10.95
North America	95.26%	83.85	52.84
Northeast Asia	95.88%	55.83	31.84
Oceania	97.62%	70.37	43.99
South Africa	18.36%	4.83	3.06
South America	94.9%	79.69	45.58
South Asia	98.71%	57.87	35.16
Southeast Asia	81.05%	33.03	13.19
Southwest Asia	79.38%	32.6	12.12
West Africa	97.95%	74.52	51.06
West Indies	89.23%	44.6	23.21
World	96.81%	68.99	39.46
Average	87.86	53.27	30.82
Standard deviation	18.6	20.31	14.7

**Table 5.** Population coverage of ZIKVac. The coverage estimation was made using Population Coverage tool available on IEDB online. <sup>a</sup>Projected population coverage. <sup>b</sup>Average number of epitope hits / HLA combinations recognized by the population. <sup>c</sup>Minimum number of epitope hits / HLA combinations recognized by 90% of the population.



**Figure 7.** CD4+ and CD8+ T cell responses to ZIKV peptides. PBMC of 10 individuals with previous infection to ZIKV were stimulated for 5 h with the ZIKV peptide pool. After 1 h of stimulation, Brefeldin A was added. The cells were stained with antibodies for the surface markers CD3, CD4 and CD8 and intracellular stained with anti-IFN- $\gamma$ , IL-2 and anti-TNF- $\alpha$ . IFN- $\gamma$ , IL-2 and TNF- $\alpha$  production by CD4+ T cells (A–C) and CD8+ T cells (D–F) was evaluated by flow cytometry. The horizontal lines indicate the medians. Mann–Whitney test was performed to determine statistical significance and  $p < 0.05$  was considered significant.

Moreover, a chimeric protein-peptide structure may be more efficient in eliciting neutralizing antibodies than the peptides alone, due to the fact that it is more probable to find discontinuous epitopes in a protein structure, while peptides usually contain only linear epitopes. Our vaccine structure revealed ten discontinuous epitopes and fourteen continuous epitopes. Interestingly, the two discontinuous epitopes (1st and 3rd best discontinuous epitopes) with the best scores were located at EDIII.

Even though the main purpose of EDIII in our vaccine is to induce neutralizing antibodies, it can also be degraded by different types of proteasomes and/or by cellular enzymes and be presented to T cells after binding to proper MHC<sup>43</sup>. Salvador et al. (2019) found by IEDB NetMHCpan method epitopes for CD8 + (MMLELDPPF) and CD4 + T cells (QMAVDMQTLTPVGR) that are present in the EDIII of our vaccine<sup>44</sup>. By utilizing different methods (NetCTL1.2 and Ellipro Protrusion index), Mirza and colleagues found 4 epitopes present in EDIII, of which 3 are CD8 epitopes (CTAAFTFTK, GTVTVEVQY, STENSKMML), and 1 is a linear B cell epitope (LKGVSYSLCTAAFTFTKI)<sup>45</sup>. Furthermore, Cunha-Neto and collaborators explored a new mathematical model of structural entropy and found that KMMLELDPPF is a CD8 epitope, and interestingly this epitope is also part of EDIII<sup>46</sup>. Several epitopes derived from EDIII were found in previous reports to induce CD8 + and CD4 + T cell responses, demonstrating that the protective potentialities of ZIKVac can extrapolate our predictions<sup>45–47</sup>.

Regarding the cellular immune response of ZIKVac, we included 12 HLA class II peptides derived from all structural and non-structural proteins of ZIKV. CD4 + T cell differentiation into effector memory and terminally differentiated cells was observed in patients in the acute phase of ZIKV infection<sup>48</sup>, which suggests that a CD4 + T cell epitope-based vaccine may contribute to eliciting a cytokine inducing response. Apart from binding affinity and promiscuity, we also considered IFN- $\gamma$  induction capability in order to select our vaccine's CD4 + T cell peptides, since high levels of this cytokine during the acute phase of ZIKV infection appears to be related to viral clearance and mild disease<sup>49</sup>. Moreover, some of the selected CD4 + T cell epitopes were described previously as being also CD8 and B cell epitopes, and interestingly, AAGAWYVYV derived from NS2B, induced IFN- $\gamma$  as detected by IFN- $\gamma$  ELISPOT<sup>50</sup>.

CD8 + T cell epitopes are critical components of our vaccine construct, since most studies support the idea that these cells help control infection through viral clearance. In mice with IFN receptor-competent T cells and dendritic cells, CD8 + T cells demonstrated a protective role during ZIKV infection<sup>50</sup>. CD4 + and CD8 + T cells proliferation assessment in mice infected with ZIKV induced a strong adaptive immune response, which helped prevent infection in the brain and testes. Moreover, in ZIKV-infected pregnant mice, CD8 + T cell responses were decreased in comparison with non-pregnant mice, which suggests that the cell-mediated immunity during pregnancy may contribute to preventing virus spread to the fetus<sup>51</sup>. In accordance with our study, Cunha-Neto et al. (2017) described the epitope QEGAVHTAL (from E protein) as potentially capable of inducing CD8 + T cell response by using structural entropy-based predictions<sup>46</sup>. Furthermore, some epitopes placed in the CD8 portion of ZIKVac (SPFGGLKRLPA, MAVDMQTLTPV) were found to induce IFN- $\gamma$  response *in vitro*<sup>47,50</sup>. Due to the limited amount of CD8 epitopes and aiming to increase population coverage as much as possible, we opted to use overlapping CD8 peptides as it would make possible to include a higher number of epitopes in a shorter amino acid sequence; however, this resulted in some peptides longer than 11 amino acids. In these cases, we would expect that intracellular peptidases, such as Endoplasmic reticulum aminopeptidase 1 and 2 (ERAP1 and 2) would play a role in cleaving these epitopes in order to allow their coupling with the HLA type I binding groove<sup>43</sup>. Some of the epitopes we selected showed negative scores for some of the parameters we analyzed. However, we believe they will not bring negative outcomes to the vaccine immunogenicity, since none of them were considered inducers of the immunoregulatory cytokines IL-10 and IL-4. Prediction results showed these epitopes are also capable of activating APCs, which can be a booster for activation of innate immunity by our vaccine. Moreover, all the four epitopes were considered proinflammatory. This can indicate that even if they are not able to induce IFN, they can be inducers of other proinflammatory cytokines such as TNF- $\alpha$ .

Other *in silico* vaccine structures based on epitope prediction have been designed since the association of ZIKV with severe disease and fetus malformation<sup>52–56</sup>. Differently from our vaccine structure, some of them were not based on consensus sequences<sup>54,56</sup>. Alam et al., (2016) and Franco et al., (2017) developed vaccines based only on the Envelope protein and NS5 protein, respectively, and Pandey et al., (2018) focused on E, C prM and NS1 proteins<sup>52,53,54</sup>. Similar to our approach, in which we included at least one peptide from each ZIKV structural and non-structural protein, Prasasty et al., (2019) and Shahid et al., (2020) explored almost all the ZIKV proteins<sup>55,56</sup>. Prasasty et al., (2019) did not have the whole vaccine construct analyzed for parameters such as autoimmunity, allergenicity and stability, but only the peptides alone<sup>55</sup>. Our analysis showed that the whole protein assessment is important to determine the assembly viability, since addition, deletion or changes in the order of peptides disposition can affect vaccine's efficacy and safety. Finally, our vaccine assembly is the only one that includes a protein domain, which we believe is an important differential feature and can be advantageous when it comes to the presence of discontinuous B cell epitopes and induction of a more robust and specific antibody response.

Interestingly, some of our epitopes were also found by Pandey et al., (2018) and Prasasty et al., (2019), even though they used different methods than ours<sup>54,55</sup>. One of our CD4 + T cell epitopes (HGPIRMVLAILAFLR) and two of our CD8 + T cell epitopes (IFRNPGFALAAAAIA; VRGAKRMAVLG) are contained in CD4 + T cell epitopes selected by Pandey et al., (2018) (LGHGPIRMVLAILAFLRFTA; VENWIFRNPGFALAAAAIAWLLGSS; VRGAKRMAVLGDTAWD). Similarly, three CD8 + T cell (MVLAILAFL; YMWLGARFL; FAAGAWYVY) and two CD4 + T cell epitopes (YRIMLSVHG; LYFHRRDLR) selected by Prasasty et al., (2019) are contained in five of our CD4 epitopes (HGPIRMVLAILAFLR; YMWLGARFLFEAL; FAAGAWYVYVKTGKR; EYRIMLSVHG-SQHSG; LLYFHRRDLRLMANA). The epitope YRIMLSVHG selected was also selected by Alam et al., (2016)<sup>52</sup>.

Adjuvant use in vaccine constructs is important for immune response enhancement. We tested our vaccine with 4 different adjuvants in order to find one that could increase its antigenicity. We selected 4 TLR agonists as adjuvants because activation of these receptors leads to innate cell stimulation, cytokine production, receptor expression, dendritic cell maturation and subsequently T cell priming<sup>57</sup>. These proteins were chosen based on



other studies that showed that they presented promising results as adjuvants. For instance, flagellin is a TLR5 agonist derived from *Bacillus subtilis* subsp. *spizizenii* strain W23 (GenBank access number: ADM39502.1) retrieved from NCBI. It has been previously shown that flagellin induced robust humoral response and pro-inflammatory cytokine production when administered with a whole inactivated influenza vaccine in mice<sup>58</sup>. Heparin-binding hemagglutinin (HBH) (UniProtKB—P9WIP9) is a TLR4 agonist derived from *Mycobacterium tuberculosis* H37Rv, while 50S ribosomal protein L7/L12 (UniProtKB—P9WHE3) is also a molecule derived from *Mycobacterium tuberculosis* H37Rv and is described as a TLR4 binder. When used as adjuvant for a therapeutic cancer vaccine, HBH favored Th1 polarization and dendritic cell (DC) function<sup>59</sup>. 50S ribosomal protein L7/L12 also enhanced Th1 polarization in a DC-based tumor immunotherapy<sup>60</sup>. Finally, RS09 (APPHALS) is a short synthetic LPS peptide mimics designed using Phage Display Libraries and described as capable of binding to TLR4<sup>61</sup>. RS09 was previously used as adjuvant in an HIV-1 vaccine that showed promising results in mice<sup>62</sup>. Our autoimmunity analysis showed that the only portion of the vaccine that had more linear amino acids similarity with human proteins was present on the Flagellin adjuvant portion (MAKEMSEF). However, this peptide is not long enough to induce a CD4 type response and is not predicted to bind to any of the HLA class I molecules considered in this study (the sequence is too short to be assessed in respect to its CD4-inducing properties). Without the help of T-cells, it is not likely that this peptide will induce a B cell response<sup>63</sup>. Furthermore, flagellin has been used as adjuvant in vaccines against bacterial pathogens<sup>64,65</sup>, parasitic organisms<sup>66,67</sup>, cancer<sup>68</sup> and virus infection, such as influenza<sup>69</sup>, WNV<sup>70</sup>, DENV<sup>71</sup> and HIV<sup>72</sup>. This adjuvant was also the one that increased vaccine antigenicity score the most in relation to the other 3, indicating that it is a potent adjuvant, however safety experimental evaluation is necessary.

Even though we presented our ZIKVac in combination with 4 protein TLR agonists, its potential application is not limited to them. ZIKVac could also be expressed or synthesized without the adjuvant domain and be administered with other types of adjuvants, such as MPL. MPL is a TLR4 agonist dependent on TRIF signaling that has been used for influenza vaccination in Brazil due to its not so aggressive immunogenicity in comparison to LPS and for reducing the dose of the antigen in vaccination regimens<sup>73</sup>. ZIKVac could also be used with aluminum, which is the most frequent adjuvant in licensed vaccines, showing a great safety record<sup>74,75</sup>. The combination of both adjuvants can lead to a more balanced Th1/Th2 response and resulted as the best adjuvant scheme when compared to either adjuvant alone or MF59<sup>35</sup>. Thus, we initially plan to express the proposed vaccine in eukaryotic cells and test it as a protein adjuvanted with alum and MPL in mice. However, it would also be particularly interesting to test it as a nanoparticle encapsulated mRNA or carried by a Hepatitis B core Antigen (HBcAg)-VLP since both strategies have been used for ZIKV vaccines and showed promising data<sup>12,76</sup>.

An ideal vaccine should have global coverage. In South America, where ZIKV infection has been more prevalent<sup>77</sup>, our vaccine's peptides covered 94.9% of the population. The highest coverage was in South Asia (98.71%), and the lowest coverage was found in South Africa (18.36%). However, the coverage in North and Central Africa were 77.17% and 95.44%, respectively. In general, world coverage was 96.81%, indicating that these epitopes are promiscuous and might be recognized by various populations worldwide. This is important because the virus may spread to unexpected regions, and a universal vaccine would be crucial to stop any novel epidemics. Similarly, when we evaluated HLA coverage of our peptides using Propred1 and Propred server, we found that ZIKVac's CD4 peptides covered 100% of the HLA molecules available in the server and CD8 epitopes covered 46 of the 47 available molecules. HLA-A\*1101 (predominantly found among the Asian population) was the only molecule to which none of our peptides were predicted to bind. Experimental validation of predicted CD8 + T cell DENV and hepatitis B virus (HBV) epitopes for this molecule showed that the predicted and laboratory results did not match<sup>78</sup>. This could be explained by the fact that HLA molecules frequently observed among the Caucasian population (e.g. HLA-A\*0201) have a larger amount of data available than Asian predominant HLA molecules, and the efficacy of epitope prediction algorithms depend directly on the training data available for each HLA molecule<sup>78</sup>.

Despite the limitations of prediction servers, various studies have been successful with experimental validation of predicted epitopes. Iwai et al., (2003) tested synthetic peptides of Paracoccidiodomycosis (PCM) gp43 protein selected using the TEPITOPE algorithm and found that peptide recognition by PCM patients correlated with the promiscuity of HLA-DR binding prediction<sup>79</sup>. Similarly, Farrell et al., (2016) found that predicted promiscuous *Mycobacterium bovis* epitopes significantly induced IFN- $\gamma$  secretion from infected cattle T cells when compared to epitopes with low promiscuity<sup>80</sup>. Moreover, Ribeiro et al., (2010) developed a DNA vaccine which encodes 18 HIV subtype B CD4 + T-cell epitopes<sup>81</sup>. The DNA vaccine induced robust CD4 + and CD8 + T cell responses in mice expressing different class II HLA alleles, and 16 of the 18 coded epitopes were recognized by PBMC from HIV infected patients. These studies suggest that in silico prediction approaches can guide the selection of conserved and antigenic epitopes that can be recognized by the target population, thereby reducing the time to develop effective vaccines. In order to test whether our peptides are capable of inducing immune response, we performed a preliminary in vitro assay and evaluated the production of cytokines after stimulation of PBMCs of individuals with previous ZIKV infection with a pool of the 25 ZIKV peptides. Both CD4 + and CD8 + T cells showed significant production of IFN- $\gamma$  upon stimulation, confirming what was predicted using the IFN epitope prediction server. IL-2 production was also detected by CD8 + T cells. The results indicate that the ZIKV peptides of our in silico designed vaccine were presented to T cells in the course of natural ZIKV infection and generated memory responses that could be boosted by peptide stimulation in vitro leading to cytokine production by CD4 + and CD8 + T cells. These data support our epitope prediction strategy and reinforce the vaccinal potential of these peptides.

## Conclusions

ZIKV infection is especially concerning in pregnant women, and more elaborated approaches are needed in order to attend this population. We designed a chimeric multidomain Zika vaccine intending to cover all the protective immune responses against ZIKV, targeting specific CD4<sup>+</sup>, CD8<sup>+</sup> T cells and B cells. This *in silico* approach gave us many promising hits. *In vitro* tests showed that our ZIKV peptides can be recognized by T cells and induce immune response. We intend on later synthesizing the whole vaccine structure to validate its potential *in vivo* as well.

## Material and methods

**Brazilian strains consensus sequence.** All complete Brazilian ZIKV strains (by March of 2018) were obtained from NCBI (<https://www.ncbi.nlm.nih.gov/>) and downloaded in FASTA protein format. The sequences access numbers are listed on Supplementary 1. The alignment was done by the software ClustalX2.1, and the consensus sequence was obtained with Jalview (version 2.11). This sequence was later separated into single proteins, which were used as input in FASTA format for the selection of potential HLA type I and II immunogenic epitopes. After aligning the sequences with ClustalX, we built a phylogenetic tree using FigTree v1.4.4.

**HLA class I-restricted epitope prediction.** In order to find the best potential ZIKV type I epitopes for immunization purposes, three binding peptides prediction platforms were compared: Propred1 (<http://crdd.osdd.net/raghava/propred1/>), IEDB Class I Immunogenicity (<http://tools.iedb.org/immunogenicity/>) and IEDB MHC-I Binding Predictions (<http://tools.iedb.org/mhci/>), following this order. Epitopes that were found promising in all 3 platforms were considered for further investigation. Propred1 highlights antigenic epitopes in a given protein sequence in 47 MHC alleles including HLA and H2 alleles<sup>82</sup>. All available alleles were selected for maximum population coverage, and the threshold of 3% was selected for the immunoproteasome and the proteasome filters. IEDB MHC-I Binding Predictions combines information from several binding prediction algorithms. The IEDB recommended method was used as it includes the Artificial Neural Network (ANN), Stabilized Matrix Method (SMM) and Combinatorial Library methods. The IEDB server has a list of the 27 most frequent class I alleles available (97% population coverage)<sup>83</sup>, which were used for the prediction considering 9mer peptides. ANN and SMM indicate the predicted binding affinity between HLA class I and a given epitope in nM, so low values (< 500 nM) of ANN and SMM indicate that the peptide and the HLA molecule may have a stable interaction. Finally, to measure peptide immunogenicity, IEDB Class I Immunogenicity resource was used. Peptides with immunogenicity score (> 0.15) were considered as immunogenic. In this tool, the amino acids properties as well as their position inside a given peptide are used to calculate the immunogenicity score of a given epitope<sup>84</sup>.

**HLA class II-restricted epitope prediction.** In order to select the best potential ZIKV CD4<sup>+</sup> T epitopes, the binding affinity to HLA class II molecules, promiscuity, immunogenicity and predicted IFN- $\gamma$  induction were evaluated. Epitopes found promising in all four parameters were selected for further analysis. Initially, CD4<sup>+</sup> T cell epitopes for ZIKV structural and non-structural proteins were predicted using Propred MHC-II Binding Peptide Prediction Server (<http://crdd.osdd.net/raghava/propred/>). 15-mer length epitopes were predicted for all 51 HLA class II alleles available in the server with 1% threshold. Propred server is based on matrices for the alleles obtained from a pocket profile database previously described<sup>85</sup>. Epitopes were selected considering the maximum number of alleles that bind to the epitope. The IEDB server (<http://tools.iedb.org/mhcii/>) was then used to determine epitope affinity to HLA class II alleles based on percentile rank and IC50 values. Percentile rank  $\leq 20$  was considered for selecting promiscuous peptides and IC50  $\leq 500$  nM was considered for peptides with intermediate to high binding affinity. Epitope immunogenicity was evaluated using the IEDB server (<http://tools.iedb.org/CD4episcore/>), considering percentile rank lower than 66 for immunogenic peptides<sup>86</sup>. Finally, epitope potential to induce IFN- $\gamma$  was evaluated using IFNepitope server (<http://crdd.osdd.net/raghava/ifnepitope/>), and peptides with positive values were considered to be capable of inducing IFN- $\gamma$  production.

**Prediction of induction of IL-10, IL-4, proinflammatory response and APC activation.** The capacity to induce IL-10 and IL-4, activate APCs, and induce proinflammatory response was evaluated for epitopes that presented negative immunogenicity scores in order to check if they would bring a negative outcome to the overall vaccine immunogenicity. Prediction of induction of IL-10 and IL-4 were performed with the downloaded IL10pred platform<sup>87</sup> and with the IL-4Pred online server (<https://webs.iitd.edu.in/raghava/il4pred/index.php>)<sup>88</sup>, respectively. The capacity to activate APCs and induce proinflammatory response was performed with the VaxinPAD online server (<https://webs.iitd.edu.in/raghava/vaxinpad/index.php>)<sup>89</sup> and Pro-Inflam online server (<http://metabiosys.iiserb.ac.in/proinflamm/index.html>)<sup>90</sup>, respectively. A 0.5 threshold was used for all these parameters as previously described<sup>91</sup>.

**Vaccine construction.** The selected CD4<sup>+</sup> and CD8<sup>+</sup> T cell epitopes were chosen from a set of peptides predicted by the aforementioned servers considering population coverage, immunogenicity score and MHC binding affinity. ZIKV Envelope E Domain III (EDIII) (consensus) was added to the structure as well. The epitopes and EDIII were flanked by linkers, short amino acid sequences, which have been described to be cleaved by proteasomes and lysosomes so the epitopes can bind to HLA class I and HLA class II, respectively<sup>92</sup>. In order to find a sequence of epitopes that could achieve the highest antigenicity while maintaining protein stability, we evaluated the best order in which the epitopes could be organized before choosing the final construct, such as EDIII-CD4-CD8, EDIII-CD8-CD4 and CD8-CD4-EDIII.

Aiming to increase vaccine antigenicity, its structure was evaluated with different combination of protein adjuvants, such as flagellin, Heparin-binding hemagglutinin (HBH), 50S ribosomal protein L7/L12 (50S rib) and RS09, a synthetic LPS peptide mimics designed with Phage Display Libraries<sup>61</sup>. Antigenicity, stability, safety and some physicochemical parameters of all of them were predicted using different platforms described in the following methods.

To separate the epitopes, protein domain and adjuvant, proper linkers were used. DNTAN linkers were used to separate the adjuvant from the vaccine epitopes, and/or the EDIII from the other parts of the vaccine<sup>93</sup> while GPGPG linker was used to separate CD4 epitopes<sup>94</sup>, and AAY was used to individualize CD8 epitopes<sup>92,95</sup>. We chose GPGPG linkers considering it induced broad MHC class II-restricted T cell responses in mice vaccinated with an epitope-based HIV vaccine previously designed by our group<sup>81</sup>. The DNTAN linker was chosen to give flexibility to the different parts of the vaccine construct because it is considered a loop in protein (LiP)<sup>96</sup>, which is important to mediate biological processes such as signaling and protein–ligand binding<sup>97</sup>.

**Population coverage.** Assessment of vaccine epitopes coverage is important to predict which population or populations are going to benefit from the selected epitopes. To this end, we used the IEDB online server Population Coverage (<http://tools.iedb.org/population/>) developed by Bui et al., (2006)<sup>98</sup>, utilizing the Allele Frequencies Net Database that contains a study of 1081 populations and 1066 alleles of the HLA gene<sup>99–101</sup>. The HLA class I and class II restricted epitopes category was selected, and the 12 CD4+ and 13 CD8+ T cell epitopes were included. All world regions were selected to assess all possible areas that could be covered by the vaccine.

**Vaccine safety.** If not evaluated properly, vaccines can become allergens and/or induce autoimmunity. In order to assess if the proposed vaccine could be a potential allergen, the protein FASTA sequences were used as input in the allergenicity prediction server Allergen FP v.10 (<http://ddg-pharmfac.net/AllergenFP/index.html>). Allergen FP v.1.0 method is based on four steps algorithm considering the physicochemical properties of individual amino acids in a given protein. It applies a universal fingerprint approach that is alignment-free and has shown accuracy of 88%<sup>102</sup>.

Additionally, if self-antigens are found in the vaccine sequence, the humoral and cellular responses induced against the vaccine may target self-proteins, leading to autoimmunity. In order to increase the chances that ZIKVac would not present such complication, blastp was performed with the NCBI platform (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>), using the parameters described elsewhere<sup>103</sup>. Identical continuous human sequences smaller than 6 amino acids were discarded as the likelihood of such pairing is high, but not significant<sup>103</sup>.

**Antigenicity.** Antigenicity of the final vaccine construct was predicted using VaxiJen v2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>), with a threshold of 0.4 and “virus” as target organism or with ANTIGENpro (<http://scratch.proteomics.ics.uci.edu>).

**Physicochemical properties and solubility.** Evaluation of physicochemical characteristics of proteins are crucial for proper protein expression, separation, purification, storage and additional analyses. Thus, the vaccine's physicochemical properties were analyzed with ProtParam ExPASy Server<sup>104</sup>, <https://web.expasy.org/protparam/>. With the ProtParam tool, it was observed the protein molecular weight, theoretical pI, extinction coefficients, estimated half-life, instability index, aliphatic index and GRAVY (Grand average of hydrophaticity).

**Secondary and tertiary structure.** Tertiary and secondary structures of all constructs were predicted using RaptorX Structure Prediction (<http://raptorx.uchicago.edu/StructurePrediction/predict/>). This server was chosen because it is capable of predicting tertiary structures of proteins that have less than 30% similarity with proteins deposited in PDB (Protein Data Bank), i.e. without any similar homologs; additionally, it also provides information related to the protein secondary structure, solvent accessibility and a quality score, which can be used to select the best model. RaptorX performance in CASP9 (Critical Assessment of protein Structure Prediction 9) was superior than the other participating servers when complex targets were evaluated<sup>105</sup>.

**Tertiary structure refinement and validation.** Tertiary structure refinement was conducted with 3D Refine web server (<http://sysbio.rnet.missouri.edu/3Drefine/>). The PDB file obtained with RaptorX was used as input. After complete running, the server provides 5 refined models with values of 3D refine score, GDT-TS GDT-HA, RMSD (root-mean-square-deviation), MolProbity, RWPlus. The 3D refine and RWplus scores are the model potential energy after refinement; GDT-TS and GDT-HA indicate how conservative was the refinement; RMSD indicates in angstroms the deviation score of the model; MolProbity indicates how physically realistic the refined model is.

3D Refine algorithm combines 2 steps of minimization that includes optimizing Hydrogen Bonding Network and Energy Minimization. The algorithm refinement capacity at local and global measurement of structural quality has been shown in CASP8, CASP9 and CASP10 with superior results compared to other refinement methods<sup>106</sup>.

After refinement, protein structure was validated by RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>), which provides the Ramachandran plot. The Ramachandran plot will present if the  $\Psi$  and  $\Phi$  angle clusters are well positioned and points out if there is any abnormality resulted from the protein refinement. The torsion angles  $\Psi$  and  $\Phi$  characterizes the spatial conformations of each amino acid residue, and the Ramachandran plot groups information about the number of residues in favored region, allowed region or outlier region<sup>107</sup>.

Based on the refinement scores and Ramachandran favored and disfavored regions, the best models were chosen for representation and discontinuous epitope prediction. The selected models were edited using Chimera v1.13.1.

**Continuous and discontinuous B cell epitopes.** Elucidation of conformational vaccine epitopes are important as BCR recognizes epitopes without presentation or previous processing. In addition, protein folding may hide linear B-cell epitopes. Tertiary structure of the protein was used as input in the ElliPro server (<http://tools.ieadb.org/elliPro/>), and minimum score was set as 0.5 (default) and maximum distance was 6 Å (default). ElliPro shows as output potential sequences of discontinuous epitopes as well as their location in the tertiary protein structure. This prediction is based on Thornton's method<sup>108</sup> and also includes residue-clustering algorithm. It is taken into consideration that, in case the protein is considered an ellipsoid, sequences that protrude from the ellipsoid are more freely recognized by antibodies<sup>109</sup>. ElliPro has the best AUC, 0.732, when compared with six other structure-based methods after testing of discontinuous protein epitopes from a benchmark dataset<sup>110</sup>.

**ZIKV peptide synthesis.** All 13 CD8 and 12 CD4 ZIKV peptides predicted in this work were synthesized through solid phase chromatography using 9-fluorophenylmethoxy-carbonyl strategy, with the C-terminal carboxyl group in amide form. Peptide purity and quality were done by reverse-phase high performance liquid chromatography by Biotech/Pepmic Co. (Brasília, Brazil/China), considering the peptides with purity > 95%. Peptides were dissolved in dimethylsulfoxide (DMSO). All 25 peptides were pooled together and after testing different concentrations, 25 µg/mL was determined as the best concentration for usage.

**Human blood samples collection and ZIKV infection detection.** In order to validate the immunogenicity of the selected peptides, blood samples from 10 convalescent donors previously infected with ZIKV were used to carry out in vitro assays (Supplementary 2). ZIKV infection was confirmed by plasma and/or urine detection of ZIKV RNA by RT-qPCR during the acute phase of infection. Blood collection was conducted between two and three years after the acute phase of infection diagnosis at Federal University of Goiás (Goiânia, Brazil). IgM and IgG serology for ZIKV was performed after blood collection using the double-path TR DPP<sup>®</sup> Zika IgM/IgG immunochromatographic test at the Bio-Manguinhos Recombinant Technology Laboratory in Oswaldo Cruz Foundation (Rio de Janeiro, Brazil). Peripheral Blood Mononuclear cells (PBMCs) were isolated by standard density gradient centrifugation on Ficoll-Paque™ Plus (GE Healthcare) and cryopreserved in a 90% fetal bovine serum (FBS) (GIBCO, Grand Island, USA) and 10% DMSO (Sigma-Aldrich, St Louis, USA) solution. Cells were stored in liquid nitrogen until usage.

**PBMC stimulation with ZIKV peptide pool, surface phenotype and intracellular cytokine staining.** PBMCs from patients previously infected with ZIKV were thawed and incubated at 37 °C with 5% CO<sub>2</sub> for 2 h. Cells were then washed and distributed at a concentration of 2 × 10<sup>6</sup> cells/mL in polypropylene tubes with RPMI 1640 medium (Lonza, Walkersville, USA) supplemented with 0.1% penicillin, 0.1% streptomycin (GIBCO, Grand Island, USA) and 5% FBS. Cells were stimulated for 5 h with 25 µg/mL of the ZIKV peptide pool containing all 25 CD4 and CD8 peptides at 37 °C with 5% CO<sub>2</sub>. After 1 h of stimulation, 1 µg/mL of Brefeldin A was added (BD Biosciences). DMSO was added to non-stimulated cells as negative control and Staphylococcal enterotoxin B (SEB) was used as positive control. Cells were washed with PBS and then with FACS buffer and stained with surface antibodies (CD3-APC-Cy7 (clone SK7, BD Pharmingen); CD4-PerCP-Cy5.5 (clone RPA-T4, BD Pharmingen); CD8-PE (clone HIT8a, BD Pharmingen). Cells were incubated at 4 °C for 30 min and then washed twice in FACS buffer. 100 µl of Cytofix/Cytoperm (BD Biosciences, San Diego, USA) was added and cells were incubated at 4 °C for 20 min. Cells were then washed with Perm/Wash 1X (BD Biosciences, San Diego, USA) and stained with intracellular cytokine antibodies (IFN-γ-FITC (clone B27, BD Pharmingen); IL-2-APC (clone 5344.111, BD FastImmune); TNF-α-PE-Cy7 (clone MAb11, BD Pharmingen) at 4°C for 30 min. Cells were washed twice with Perm/Wash 1X and fixed in 2% paraformaldehyde solution (Sigma-Aldrich, St. Louis, USA). Acquisition was performed on a BD FACS Canto II flow cytometer (BD Biosciences) and a minimum of 400,000 events were recorded from each tube. Data were analyzed using FlowJo software (version 10.0, Tree Stat, Ashland, OR, USA).

**Data analysis.** Kolmogorov–Smirnov and Shapiro–Wilk were used to test the normality of the data. The non-parametric Mann–Whitney test was used to analyze differences between non-stimulated and peptide stimulated conditions. P values < 0.05 were considered statistically significant. Statistical analysis was performed using Prism version 10 (GraphPad Prism Software Inc., San Diego, CA, USA).

**Ethics statement.** Ethical approval for this study was obtained from the Institutional Review Board of Pontifical Catholic University of Goiás (CEP—Research Ethics Committee) under the protocol number 46073815.9.0000.00370. All patients signed a written informed consent form before the interview and sample collection. The authors also declare that all methods were carried out in accordance with relevant guidelines and regulations.



## References

- Pan American Health Organization. Zika. *Zika* (2019).
- World Health Organization. ZIKA EPIDEMIOLOGY UPDATE July 2019. 1–14 (2019).
- Hills, S. L., Fischer, M. & Petersen, L. R. Epidemiology of Zika virus infection. *J. Infect. Dis.* **216**, S868–S874 (2017).
- Pan American Health Organization. *Zika-Epidemiological Report*. (2017).
- Hickman, H. D. & Pierson, T. C. T cells take on Zika virus. *Immunity* **46**, 13–14 (2017).
- Goncé, A. *et al.* Spontaneous abortion associated with Zika virus infection and persistent viremia. *Emerg. Infect. Dis.* **24**, 933–935 (2018).
- Mlakar, J. *et al.* Zika Virus associated with microcephaly. *N. Engl. J. Med.* **374**, 951–958 (2016).
- Weaver, S. C., Charlier, C., Vasilakis, N. & Lecuit, M. Zika, chikungunya, and other emerging vector-borne viral diseases. *Annu. Rev. Med.* **69**, 395–408 (2018).
- Dawes, B. E. *et al.* Research and development of Zika virus vaccines. *Npj Vaccines* **1**, 16007 (2016).
- Robbiani, D. F. *et al.* Recurrent potent human neutralizing antibodies to Zika virus in Brazil and Mexico. *Cell* **169**, 597–609.e11 (2017).
- Yang, M., Dent, M., Lai, H., Sun, H. & Chen, Q. Immunization of Zika virus envelope protein domain III induces specific and neutralizing immune responses against Zika virus. *Vaccine* **35**, 4287–4294 (2017).
- Yang, M., Lai, H., Sun, H. & Chen, Q. Virus-like particles that display Zika virus envelope protein domain III induce potent neutralizing immune responses in mice. *Sci. Rep.* **7**, 1–12 (2017).
- Lucas, C. G. O. *et al.* Critical role of CD4+ T cells and IFN $\gamma$  signaling in antibody-mediated resistance to Zika virus infection. *Nat. Commun.* **9**, 3136 (2018).
- Koblischke, M. *et al.* Structural influence on the dominance of virus-specific CD4 T cell epitopes in Zika virus infection. *Front. Immunol.* **9**, (2018).
- Elong Ngono, A. & Shresta, S. Cross-reactive T cell immunity to dengue and Zika viruses: New insights into vaccine development. *Front. Immunol.* **10**, 1316 (2019).
- Hassert, M. *et al.* CD4+T cells mediate protection against Zika associated severe disease in a mouse model of infection. *PLoS Pathog.* **14**, e1007237 (2018).
- Elong Ngono, A. *et al.* Mapping and role of the CD8+ T cell response during primary Zika virus infection in mice. *Cell Host Microbe* **21**, 35–46 (2017).
- Huang, H. *et al.* CD8+ T cell immune response in immunocompetent mice during Zika Virus infection. *J. Virol.* **91**, 1–15 (2017).
- El Sahly, H. M. *et al.* Clinical, virologic, and immunologic characteristics of Zika virus infection in a cohort of US patients: Prolonged RNA detection in whole blood. *Open Forum Infect. Dis.* **6**, 1–9 (2019).
- Savino, W. *et al.* Zika virus infection in the elderly: Possible relationship with guillain-barré syndrome. *Gerontology* **63**, 210–215 (2017).
- Schwartzmann, P. V. *et al.* Zika virus meningoencephalitis in an immunocompromised patient. *Mayo Clin. Proc.* **92**, 460–466 (2017).
- Poland, G. A., Ovsyannikova, I. G. & Kennedy, R. B. Zika vaccine development: Current status. *Mayo Clin. Proc.* **94**, 2572–2586 (2019).
- Nambala, P. & Su, W. C. Role of Zika virus prM protein in viral pathogenicity and use in vaccine development. *Front. Microbiol.* **9**, 1–6 (2018).
- Saiz, J.-C. Therapeutic advances against ZIKV: A quick response, a long way to go. *Pharmaceuticals* **12**, 127 (2019).
- Zhang, L. Multi-epitope vaccines: A promising strategy against tumors and viral infections. *Cell. Mol. Immunol.* **15**, 182–184 (2018).
- Dawood, R. M. *et al.* A multiepitope peptide vaccine against HCV stimulates neutralizing humoral and persistent cellular responses in mice. *BMC Infect. Dis.* **19**, 1–11 (2019).
- Gilbert, P. B. *et al.* Long-term safety analysis of preventive HIV-1 vaccines evaluated in AIDS vaccine evaluation group NIAID-sponsored Phase I and II clinical trials. *Vaccine* **21**, 2933–2947 (2003).
- Keskin, D. B. *et al.* Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial. *Nature* **565**, 234–239 (2019).
- Depoil, D. *et al.* Early events of B cell activation by antigen. *Sci. Signal.* **2**, pt1 (2009).
- Costa, L. C. *et al.* New insights on the Zika virus arrival in the Americas and spatiotemporal reconstruction of the epidemic dynamics in Brazil. *Viruses* **13**, 12 (2020).
- Reynolds, M. R. *et al.* Vital Signs: Update on Zika Virus—Associated birth defects and evaluation of all US infants with congenital Zika Virus exposure. *Morb. Mortal. Wkly. Rep.* **66**, 366–373 (2017).
- Shapiro-Mendoza, C. K. *et al.* Pregnancy outcomes after maternal Zika Virus infection during pregnancy—US territories, January 1, 2016–April 25, 2017. *MMWR. Morb. Mortal. Wkly. Rep.* **66**, 615–621 (2017).
- Morrey, J. D. *et al.* Humanized monoclonal antibody against west Nile virus envelope protein administered after neuronal infection protects against lethal encephalitis in hamsters. *J. Infect. Dis.* **194**, 1300–1308 (2006).
- Lai, P.-Y. *et al.* Production of a neutralizing antibody against envelope protein of dengue virus type 2 using the linear array epitope technique. *J. Gen. Virol.* **95**, 2155–2165 (2014).
- Wang, X. *et al.* Effects of adjuvants on the immunogenicity and efficacy of a Zika Virus envelope domain III subunit vaccine. *Vaccines* **7**, 161 (2019).
- López-Camacho, C. *et al.* Immunogenicity and efficacy of Zika virus envelope domain III in DNA, protein, and ChAdOx1 adenoviral-vectored vaccines. *Vaccines* **8**, 1–20 (2020).
- Tai, W. *et al.* Critical neutralizing fragment of Zika virus EDIII elicits cross-neutralization and protection against divergent Zika viruses article. *Emerg. Microbes Infect.* **7**, 7 (2018).
- Beasley, D. W. C. & Barrett, A. D. T. Identification of neutralizing epitopes within structural domain III of the West Nile virus envelope protein. *J. Virol.* **76**, 13097–13100 (2002).
- Maillard, R. A., Jordan, M., Beasley, D. W. C., Barrett, A. D. T. & Lee, J. C. Long range communication in the envelope protein domain III and its effect on the resistance of west Nile virus to antibody-mediated neutralization. *J. Biol. Chem.* **283**, 613–622 (2008).
- Chiang, C.-Y. *et al.* Lipidated dengue-2 envelope protein domain III independently stimulates long-lasting neutralizing antibodies and reduces the risk of antibody-dependent enhancement. *PLoS Negl. Trop. Dis.* **7**, 2432 (2013).
- Stettler, K. *et al.* Specificity, cross-reactivity, and function of antibodies elicited by Zika virus infection. *Science* **353**, 823–826 (2016).
- Wu, Y. *et al.* Neutralization of Zika virus by germline-like human monoclonal antibodies targeting cryptic epitopes on envelope domain III. *Emerg. Microbes Infect.* **6**, 89 (2017).

43. Lázaro, S., Gamarra, D. & Del Val, M. Proteolytic enzymes involved in MHC class I antigen processing: A guerrilla army that partners with the proteasome. *Mol. Immunol.* **68**, 72–76 (2015).
44. Salvador, E. A., Pires de Souza, G. A., Cotta Malaquias, L. C., Wang, T. & Leomil Coelho, L. F. Identification of relevant regions on structural and nonstructural proteins of Zika virus for vaccine and diagnostic test development: An in silico approach. *New Microbes New Infect.* **29**, 100506 (2019).
45. Usman Mirza, M. *et al.* Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. *Sci. Rep.* **6**, 1–17 (2016).
46. Cunha-Neto, E. *et al.* An approach for a synthetic CTL vaccine design against Zika flavivirus using class I and class II epitopes identified by computer modeling. *Front. Immunol.* **8**, 1–12 (2017).
47. Wen, J. *et al.* Identification of Zika virus epitopes reveals immunodominant and protective roles for dengue virus cross-reactive CD8+ T cells. *Nat. Microbiol.* **2**, 17036 (2017).
48. Cimini, E. *et al.* Human Zika infection induces a reduction of IFN- $\gamma$  producing CD4 T-cells and a parallel expansion of effector V $\delta$ 2 T-cells. *Sci. Rep.* **7**, 6313 (2017).
49. de Barros, J. B. *et al.* Acute Zika virus infection in an endemic area shows modest proinflammatory systemic immunoactivation and cytokine-symptom associations. *Front. Immunol.* **9**, 821 (2018).
50. Elong Ngonu, A. *et al.* Mapping and role of the CD8+T cell response during primary zika virus infection in mice. *Cell Host Microbe* **21**, 35–46 (2017).
51. Winkler, C. W. *et al.* Adaptive immune responses to zika virus are important for controlling virus infection and preventing infection in brain and testes. *J. Immunol.* **198**, 1601949 (2017).
52. Alam, A., Ali, S., Ahamad, S., Malik, M. Z. & Ishrat, R. From ZikV genome to vaccine: in silico approach for the epitope-based peptide vaccine against Zika virus envelope glycoprotein. *Immunology* **149**, 386–399 (2016).
53. Franco, L., dos, S., Oliveira Vidal, P. & Amorim, J. H. In silico design of a Zika virus non-structural protein 5 aiming vaccine protection against zika and dengue in different human populations. *J. Biomed. Sci.* **24**, 88 (2017).
54. Kumar Pandey, R., Ojha, R., Mishra, A. & Kumar Prajapati, V. Designing B- and T-cell multi-epitope based subunit vaccine using immunoinformatics approach to control Zika virus infection. *J. Cell. Biochem.* **119**, 7631–7642 (2018).
55. Prasasty, V. D. *et al.* Peptide-based subunit vaccine design of T-and b-cells multi-epitopes against zika virus using immunoinformatics approaches. *Microorganisms* **7**, 226 (2019).
56. Shahid, F., Ashfaq, U. A., Javaid, A. & Khalid, H. Immunoinformatics guided rational design of a next generation multi epitope based peptide (MEBP) vaccine by exploring Zika virus proteome. *Infect. Genet. Evol.* **80**, 104199 (2020).
57. Gnjatic, S., Sawhney, N. B. & Bhardwaj, N. Toll-like receptor agonists: Are they good adjuvants?. *Cancer J.* **16**, 382–391 (2010).
58. Skountzou, I. *et al.* Salmonella flagellins are potent adjuvants for intranasally administered whole inactivated influenza vaccine. *Vaccine* **28**, 4103–4112 (2010).
59. Jung, I. D. *et al.* Enhanced efficacy of therapeutic cancer vaccines produced by co-treatment with Mycobacterium tuberculosis heparin-binding hemagglutinin, a novel TLR4 agonist. *Cancer Res.* **71**, 2858–2870 (2011).
60. Lee, S. J. *et al.* A potential protein adjuvant derived from Mycobacterium tuberculosis Rv0652 enhances dendritic cells-based tumor immunotherapy. *PLoS ONE* **9**, 1–11 (2014).
61. Shanmugam, A. *et al.* Synthetic toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants. *PLoS One* **7**, e30839 (2012).
62. Li, M., Jiang, Y., Gong, T., Zhang, Z. & Sun, X. Intranasal Vaccination against HIV-1 with Adenoviral Vector-Based Nanocomplex Using Synthetic TLR-4 Agonist Peptide as Adjuvant. *Mol. Pharm.* **13**, 885–894 (2016).
63. Amorim, K. N. S. *et al.* The presence of T cell epitopes is important for induction of antibody responses against antigens directed to DEC205+ dendritic cells. *Sci. Rep.* **6**, 39250 (2016).
64. Lee, L. H. *et al.* Evaluation of a truncated recombinant flagellin subunit vaccine against *Campylobacter jejuni*. *Infect. Immun.* **67**, 5799–5805 (1999).
65. Rady, H. F., Dai, G., Huang, W., Shellito, J. E. & Ramsay, A. J. Flagellin encoded in gene-based vector vaccines is a route-dependent immune adjuvant. *PLoS ONE* **11**, e0148701 (2016).
66. Bargieri, D. Y. *et al.* New malaria vaccine candidates based on the Plasmodium vivax Merozoite Surface Protein-1 and the TLR-5 agonist Salmonella Typhimurium FliC flagellin. *Vaccine* **26**, 6132–6142 (2008).
67. Yin, G. *et al.* An Eimeria vaccine candidate based on Eimeria tenella immune mapped protein 1 and the TLR-5 agonist Salmonella typhimurium FliC flagellin. *Biochem. Biophys. Res. Commun.* **440**, 437–442 (2013).
68. Zheng, J. H. *et al.* Two-step enhanced cancer immunotherapy with engineered *Salmonella typhimurium* secreting heterologous flagellin. *Sci. Transl. Med.* **9**, eaak9537 (2017).
69. Turley, C. B. *et al.* Safety and immunogenicity of a recombinant M2e-flagellin influenza vaccine (STF2.4xM2e) in healthy adults. *Vaccine* **29**, 5145–5152 (2011).
70. McDonald, W. F. *et al.* A west nile virus recombinant protein vaccine that coactivates innate and adaptive immunity. *J. Infect. Dis.* **195**, 1607–1617 (2007).
71. Liu, G. *et al.* Immunogenicity and efficacy of flagellin-envelope fusion dengue vaccines in mice and monkeys. *Clin. Vaccine Immunol.* **22**, 516–525 (2015).
72. Vassilieva, E. V. *et al.* Enhanced mucosal immune responses to HIV virus-like particles containing a membrane-anchored adjuvant. *MBio* **2**, e00328 (2011).
73. Precioso, A. R. *et al.* A phase I randomized, double-blind, controlled trial of 2009 influenza A (H1N1) inactivated monovalent vaccines with different adjuvant systems. *Vaccine* **29**, 8974–8981 (2011).
74. Marrack, P., McKee, A. S. & Munks, M. W. Towards an understanding of the adjuvant action of aluminium. *Nat. Rev. Immunol.* **9**, 287–293 (2009).
75. Shah, R. R., Hassett, K. J. & Brito, L. A. Overview of vaccine adjuvants: Introduction, history, and current status. *Methods Mol. Biol.* **1494**, 1–13 (2017).
76. Richner, J. M. *et al.* Modified mRNA vaccines protect against Zika virus infection. *Cell* **168**, 1114–1125.e10 (2017).
77. Rabaan, A. A., Bazzi, A. M., Al-Ahmed, S. H., Al-Ghath, M. H. & Al-Tawfiq, J. A. Overview of Zika infection, epidemiology, transmission and control measures. *J. Infect. Public Health* **10**, 141–149 (2017).
78. Rivino, L. *et al.* Defining CD8+ T cell determinants during human viral infection in populations of Asian ethnicity. *J. Immunol.* **191**, 4010–4019 (2013).
79. Iwai, L. K. *et al.* In silico prediction of peptides binding to multiple HLA-DR molecules accurately identifies immunodominant epitopes from gp43 of Paracoccidioides brasiliensis frequently recognized in primary peripheral blood mononuclear cell responses from sensitized ind. *Mol. Med.* **9**, 209–219 (2003).
80. Farrell, D. *et al.* Integrated computational prediction and experimental validation identifies promiscuous T cell epitopes in the proteome of Mycobacterium bovis. *Microb. Genom.* **2**, e000071 (2016).
81. Ribeiro, S. P. *et al.* A vaccine encoding conserved promiscuous HIV CD4 epitopes induces broad T cell responses in mice transgenic to multiple common HLA class II molecules. *PLoS ONE* **5**, 1–9 (2010).
82. Singh, H. & Raghava, G. P. S. ProPred1: Prediction of promiscuous MHC class-I binding sites. *Bioinformatics* **19**, 1009–1014 (2003).

83. Weiskopf, D. *et al.* Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E2046–E2053 (2013).
84. Calis, J. J. A. *et al.* Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput. Biol.* **9**, e1003266 (2013).
85. Sturniolo, T. *et al.* Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat. Biotechnol.* **17**, 555–561 (1999).
86. Dhanda, S. K. *et al.* Predicting HLA CD4 immunogenicity in human populations. *Front. Immunol.* **9**, 1–14 (2018).
87. Nagpal, G. *et al.* Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. *Sci. Rep.* **7**, 1–10 (2017).
88. Dhanda, S. K., Gupta, S., Vir, P. & Raghava, G. P. Prediction of IL4 inducing peptides. *Clin. Dev. Immunol.* **2013**, 263952 (2013).
89. Nagpal, G., Chaudhary, K., Agrawal, P. & Raghava, G. P. S. Computer-aided prediction of antigen presenting cell modulators for designing peptide-based vaccine adjuvants. *J. Transl. Med.* **16**, 1–15 (2018).
90. Gupta, S., Madhu, M. K., Sharma, A. K. & Sharma, V. K. ProInflam: A webserver for the prediction of proinflammatory antigenicity of peptides and proteins. *J. Transl. Med.* **14**, 1–10 (2016).
91. Porto, S. P. *et al.* Immunoinformatic construction of an adenovirus-based modular vaccine platform and its application in the design of a SARS-CoV-2 vaccine. *Infect. Genet. Evol.* **85**, 104489 (2020).
92. Nezafat, N., Ghasemi, Y., Javadi, G., Khoshnoud, M. J. & Omidinia, E. A novel multi-epitope peptide vaccine against cancer: An in silico approach. *J. Theor. Biol.* **349**, 121–134 (2014).
93. Arai, R., Ueda, H., Kitayama, A., Kamiya, N. & Nagamune, T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng.* **14**, 529–532 (2001).
94. Livingston, B. *et al.* A rational strategy to design multi-epitope immunogens based on multiple Th lymphocyte epitopes. *J. Immunol.* **168**, 5499–5506 (2002).
95. Dolenc, I., Seemüller, E. & Baumeister, W. Decelerated degradation of short peptides by the 20S proteasome. *FEBS Lett.* **434**, 357–361 (1998).
96. Michalsky, E., Goede, A. & Preissner, R. Loops In Proteins (LIP) - A comprehensive loop database for homology modelling. *Protein Eng.* **16**, 979–985 (2003).
97. Shehu, A. & Kaviraki, L. E. Modeling structures and motions of loops in protein molecules. *Entropy* **14**, 252–290 (2012).
98. Bui, H. H. *et al.* Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinf.* **7**, 1–5 (2006).
99. Sidney, J., Grey, H. M., Kubo, R. T. & Sette, A. Practical, biochemical and evolutionary implications of the discovery of HLA class I supermotifs. *Immunol. Today* **17**, 261–266 (1996).
100. Southwood, S. *et al.* Several common HLA-DR types share largely overlapping peptide binding repertoires. *J. Immunol.* **160**, 3363–3373 (1998).
101. Gonzalez-Galarza, F. F., Christmas, S., Middleton, D. & Jones, A. R. Allele frequency net: A database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res.* **39**, D913–D919 (2011).
102. Dimitrov, I., Naneva, L., Doytchinova, I. & Bangov, I. AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics* **30**, 846–851 (2014).
103. Tan, P. T. *et al.* Conservation and diversity of influenza A H1N1 HLA-restricted T cell epitope candidates for epitope-based vaccines. *PLoS ONE* **5**, 8754 (2010).
104. Gasteiger, E. *et al.* Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook* (ed. Walker, J.) 571–607 (Humana Press, 2005).
105. Källberg, M. *et al.* Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **7**, 1511–1522 (2012).
106. Bhattacharya, D., Nowotny, J., Cao, R. & Cheng, J. 3Drefine: An interactive web server for efficient protein structure refinement. *Nucleic Acids Res.* **44**, W406–W409 (2016).
107. Kleywegt, G. J. & Jones, T. A. Phi/Psi-chology: Ramachandran revisited. *Structure* **4**, 1395–1400 (1996).
108. Thornton, J. M., Edwards, M. S., Taylor, W. R. & Barlow, D. J. Location of ‘continuous’ antigenic determinants in the protruding regions of proteins. *EMBO J.* **5**, 409–413 (1986).
109. Mukonyora, M. A review of important discontinuous B-cell epitope prediction tools. *J. Clin. Cell. Immunol.* **06**, 1–5 (2015).
110. Ponomarenko, J. *et al.* ElliPro: A new structure-based tool for the prediction of antibody epitopes. *BMC Bioinf.* **9**, 514 (2008).

## Acknowledgements

The authors acknowledge the researchers who developed the platforms and algorithms used for this work and also the Institutions that made it available for public use. The authors also acknowledge Dr. Concepta McManus, Dr Lee Chen Chen, Dr Macks Wendhell Gonçalves and Dr Natasha Matsunaga for helping to clarify some statistical analysis doubts.

## Author contributions

S.G.F. conceived and designed the study. A.C.B.A. and V.P.A. performed the in silico predictions and data analysis. S.G.F., V.P.A. and A.C.B.A. wrote the manuscript. E.C.N., I.A.H.P., A.Q.M., M.M.B., A.L.B. and R.O.R. critically discussed the results and made improvement in the vaccine approach. J.M.S. conducted the isolation of PBMCs and serology diagnosis of ZIKV. I.A.H.P. and S.G.F. performed the RT-qPCR for ZIKV. A.C.B.A. and F.O.F.C. carried out the in vitro assays and performed the cytometry analysis. A.L.B. and R.O.R. provided financial funding for the synthesis of the peptides and antibody purchase. All authors reviewed and approved the manuscript.

## Funding

This work was funded by the Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) and Programa Pesquisa para o Sistema Único de Saúde (PPSUS) (FAPEG/SES-GO/CNPq/MS-DECIT/2017, under grant n. 201710267001260 (to SGF)., and Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF) under grant n. 0193.000417/2016 (to ROR and ALB). ACBA received a research fellowship (134145/2018–6) from Conselho Nacional de Pesquisa (CNPq) as this work was part of a Masters project.

## Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-03990-6>.

**Correspondence** and requests for materials should be addressed to S.G.F.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022