

1 **Identification of a Zika NS2B epitope for which absence of IgG response is**  
2 **associated with severe neurological symptoms and the design of a**  
3 **biomarker capable of discriminatory diagnostics between severe and non-**  
4 **severe clinical phenotypes**

5 Felix F. Loeffler<sup>a,¶</sup>, Isabelle F.T. Viana<sup>b,¶</sup>, Nico Fischer<sup>c</sup>, Danilo F. Coêlho<sup>b,d</sup>, Carolina S. Silva<sup>e</sup>, Antônio F. Purificação  
6 Jr.<sup>b</sup>, Catarina M.C.S. Araújo<sup>b</sup>, Bruno H.S. Leite<sup>b</sup>, Ricardo Durães-Carvalho<sup>f</sup>, Tereza Magalhães<sup>g</sup>, Clarice N.L. Morais<sup>b</sup>,  
7 Marli T. Cordeiro<sup>b</sup>, Roberto D. Lins<sup>2,&</sup>, Ernesto T.A. Marques<sup>b,h,&</sup> and Thomas Jaenisch<sup>c,i,&\*</sup>

8 a. Max Planck Institute of Colloids and Interfaces, Department of Biomolecular Systems, Potsdam, Germany.

9 b. Department of Virology, Aggeu Magalhães Institute, Oswaldo Cruz Foundation, Recife, PE, Brazil.

10 c. Section Clinical Tropical Medicine, Department of Infectious Diseases, Heidelberg University Hospital, Germany.

11 d. Department of Fundamental Chemistry, Federal University of Pernambuco, Recife, PE, Brazil.

12 e. Department of Chemical Engineering, Federal University of Pernambuco, Recife, PE, Brazil.

13 f. Laboratory of Virology, University of Campinas, Campinas, SP, Brazil.

14 g. Arthropod-borne and Infectious Diseases Laboratory (AIDL), Department of Microbiology, Immunology and Pathology,  
15 Colorado State University, Fort Collins, CO, United States of America.

16 h. Department of Infectious Diseases and Microbiology, University of Pittsburgh, Pittsburgh, PA, United States of America

17 i. German Centre for Infection Research (DZIF), Heidelberg Site, Heidelberg, Germany.

18 ¶ These authors contributed equally to this work.

19 & These authors also contributed equally to this work.

20 \*Corresponding author: thomas.jaenisch@urz.uni-heidelberg.de

21

22 **Abstract**

23 The identification of specific biomarkers for Zika infection and its clinical complications is  
24 fundamental to mitigate the infection spread, which has been associated with a broad range  
25 of neurological sequelae. We present the characterization of antibody responses in serum  
26 samples from individuals infected with Zika, presenting non-severe (classical) and severe  
27 (neurological disease) phenotypes, with high-density peptide arrays comprising the Zika NS1  
28 and NS2B proteins. The data pinpoints one strongly IgG-targeted NS2B epitope in non-severe  
29 infections, which is absent in Zika patients, where infection progressed to the severe  
30 phenotype. This differential IgG profile between the studied groups was confirmed by  
31 multivariate data analysis. Molecular dynamics simulations and circular dichroism have  
32 shown that the peptide in solution presents itself in a sub-optimal conformation for antibody  
33 recognition, which led us to computationally engineer an artificial protein able to stabilize  
34 the NS2B epitope structure. The engineered protein was used to interrogate paired samples  
35 from mothers and their babies presenting Zika-associated microcephaly and confirmed the  
36 absence of NS2B IgG response in those samples. These findings suggest that the assessment  
37 of antibody responses to the herein identified NS2B epitope is a strong candidate biomarker  
38 for the diagnosis and prognosis of Zika-associated neurological disease.

39

40

## 41 Introduction

42 Zika virus (ZIKV) emerged in the Americas, causing an unprecedented epidemic of microcephaly in babies born  
43 to mothers infected during pregnancy, and neurological disease in adults following acute infection. Some  
44 uncertainty still remains regarding the time of introduction into the Americas; however, the virus likely entered  
45 Brazil in 2013,<sup>1</sup> with the first cases of microcephaly reported in 2015. The temporal correlation between ZIKV  
46 introduction in Brazil and the microcephaly epidemic led the Brazilian Government to hypothesize the presence of  
47 a causal association.<sup>2, 3</sup> Shortly after (1 February 2016), the World Health Organization (WHO) declared a 'Public  
48 Health Emergency of International Concern' for the clusters of microcephaly and other neurological disorders, which  
49 was only lifted in November 2016.<sup>4</sup>

50 The causal association between congenital ZIKV infection and microcephaly (now referred to as Congenital  
51 Zika Syndrome [CZS] due to its broad range of clinical manifestations) was accepted by WHO in 2016,<sup>5, 6</sup> and evidence  
52 continued to accumulate in the following years.<sup>7</sup> In addition to that, explosive outbreaks in large populations in Latin  
53 America revealed other severe neurological sequelae of ZIKV infection in children and adults, including Guillain-  
54 Barré syndrome: an immune-mediated demyelinating motor and sensory peripheral neuropathy leading to  
55 paralysis.<sup>8</sup> The ability of ZIKV to cause neurological disease is not unique among flaviviruses.<sup>9</sup> Several candidate  
56 neurovirulence mechanisms have been postulated, among them the glycosylation of the envelope protein,<sup>10</sup> the  
57 presence of neuronal receptors (such as AXL) only recognized by ZIKV, and the mutation S139N in the precursor  
58 membrane protein of ZIKV, which enhances neurovirulence possibly by creating a new receptor for progenitor cells.<sup>8</sup>  
59 <sup>11</sup> To date, however, the viral determinants of ZIKV neurovirulence and the immune components involved have not  
60 been fully unraveled.

61 Prior infection with dengue virus (DENV) has been suggested to be associated with more severe manifestations  
62 in ZIKV infections.<sup>12</sup> In countries like Brazil, more than 90% of the adult population has been previously exposed to  
63 DENV.<sup>13</sup> ZIKV and DENV are both members of the *Flaviviridae* family and exhibit considerable cross-reactivity in  
64 serological tests, which proves the close phylogenetic and antigenic relationship between these viruses<sup>14-16</sup>. High  
65 anti-DENV titers have been reported to be linked with protection from Zika,<sup>17</sup> whereas sub-neutralizing levels of  
66 anti-DENV have been shown to enhance ZIKV infection in vitro.<sup>12, 18, 19</sup>

67 The overlapping clinical syndromes and geographical distribution make differential diagnosis a major challenge,  
68 especially in situations where the causative viral agent cannot directly be detected and diagnosis is based on indirect  
69 serological tests. Highly specific serological tests, capable of discriminating between different flaviviruses and  
70 therefore providing accurate prognosis of severe disease, are so far not available. Several serological test platform  
71 candidates are currently under development and validation;<sup>20-24</sup> however, an accurate diagnostic and prognostic  
72 marker for ZIKV infection and its severe neurological manifestations has yet to be demonstrated.<sup>25</sup>

73 We aimed at characterizing important ZIKV-specific epitopes by analyzing serum samples taken from  
74 patients with different clinical outcomes after ZIKV infection in Recife, Northeast Brazil, the epicenter of the ZIKV  
75 outbreak in the Americas. With this objective, we explored a novel high-density peptide array technology to search  
76 the whole proteome of ZIKV for serological biomarkers of infection. Several platforms based on peptide arrays have  
77 been developed and applied for infectious disease research.<sup>26, 27</sup> While most reports qualify the arrays using the  
78 detection of linear epitopes binding well-characterized monoclonal antibodies, recently, we and others have used  
79 this technology to demonstrate the readout of antibody profiles against malaria in different patient sera,<sup>28</sup> as well  
80 as in Lyme disease patients, in the identification of ZIKV epitopes<sup>29</sup> and for general vaccine studies.<sup>29-31</sup> Using our  
81 novel solid material-based combinatorial synthesis method,<sup>32, 33</sup> the whole peptidome of ZIKV was produced on a  
82 unique single glass slide as overlapping 15-mer peptides.

83 Due to the complexity and multivariate nature of the data obtained with this new peptide array technology,  
84 multivariate analysis techniques rise as alternatives to explore the dataset and to recognize consistently patterned  
85 structures that correlate groups of individuals with different clinical outcomes and specific peptide sequences. In  
86 this scenario, exploratory and classification techniques were employed to screen our peptide array against ZIKV  
87 clinical samples and identify one epitope sequence as a potential biomarker for ZIKV infection and its neurological  
88 manifestation with statistical reliability. Following peptide identification, the peptide's structural and  
89 thermodynamic stability were assessed by molecular dynamics simulations and confirmed by circular dichroism  
90 analysis. To improve the performance of the identified epitope as an immunological marker, we computationally  
91 designed a new protein that mimics the natural structure of the peptide. The potential use of this protein as a  
92 prognostic marker for ZIKV infections with neurological complications was then confirmed.

93

94

## 95 Results and discussion

### 96 Identification of an immunoreactive peptide within the ZIKV NS2B protein by high-density peptide array and Multivariate 97 Analysis

98 The customized peptide array containing 1,878 15-mer peptides covering the whole proteome of the ZIKV was  
99 first used to screen acute and convalescent ZIKV samples, stratified by the presence or absence of DENV background  
100 (Tables S1 and S2 and Figure S1), to determine the overall IgG-binding properties towards specific peptides. The  
101 samples were categorized into eight groups consisting of i) acute Zika patients without dengue infection history  
102 (n=14), ii) acute Zika patients with dengue infection history (n=17), iii) acute Zika patients with low antibody titers  
103 against dengue (n=3, this group presents negative results for Dengue PRNT and Dengue IgG capture ELISA, and was  
104 included due to the difficulty to define the comparison group of 'convalescent Zika patients without dengue infection  
105 history' since Dengue seroprevalence in Recife can be above 90%<sup>13,34</sup>), iv) convalescent Zika patients without dengue  
106 infection history (n=14), v) convalescent Zika patients with dengue infection history (n=17), vi) convalescent Zika  
107 patients with low antibody titers against dengue (n=3), vii) Zika infections with neurological symptoms (NeuroZIKV)  
108 without dengue infection history (n=11), and viii) Zika infections with neurological symptoms (NeuroZIKV) with  
109 dengue infection history (n=24). Important to note that all of the NeuroZIKV serum samples used in the study have  
110 exhibited ZIKV IgG positive results in previous tests (IgG ELISA and PRNT). Serum samples from individuals not  
111 exposed to ZIKV or DENV were used as assay control group (n=8).

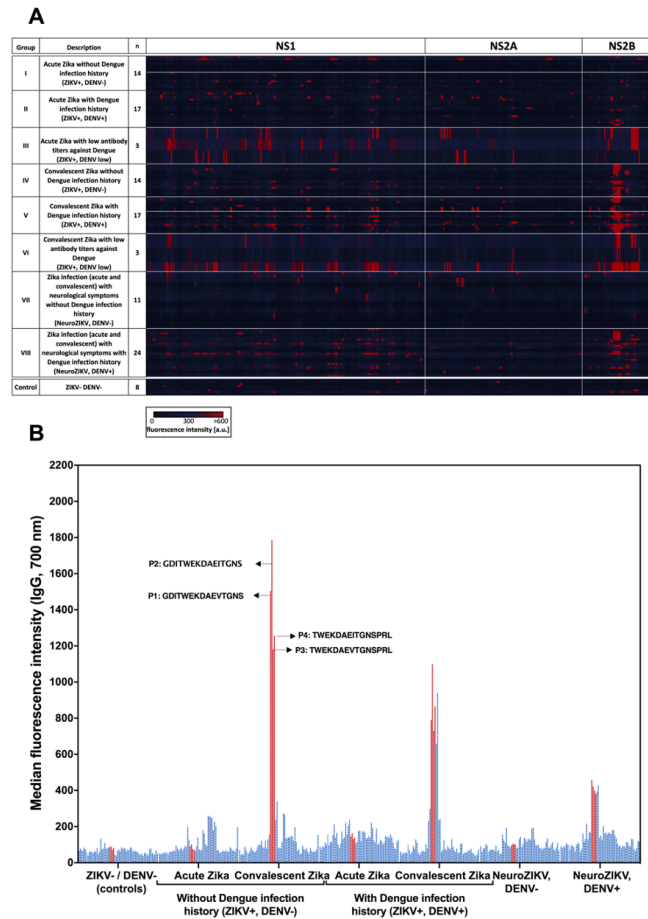
112 The median fluorescence intensity data obtained for each group showed specific IgG responses from patients  
113 with confirmed ZIKV infection in convalescent *versus* acute samples towards several peptides of the NS1 and NS2B  
114 proteins (for clarity purposes, only the data from NS1 and NS2 proteins is shown in Figure 1).

115 A second peptide array, composed of 401 15-mer peptides only covering the amino acid sequences of NS1, NS2A  
116 and NS2B proteins, was used to confirm these findings. The fluorescence intensity data of both arrays is shown in  
117 Figure 1A. Individual analysis of the median for each NS2B peptide revealed four strongly immunoreactive amino  
118 acid sequences, termed P1 (GDITWEKDAEVTGNS), P2 (GDITWEKDAEITGNS), P3 (TWEKDAEVTGNSPRL) and P4  
119 (TWEKDAEITGNSPRL) (Figure 1B). The corresponding peptides are adjacent; consequently, their sequences are  
120 overlapping. Combining the two sequences led to the identification of a major NS2B epitope  
121 (GDITWEKDAEV(I)TGNSPRLDVA) located at the ZIKV poly-proteome position (1425–1445). Figure 1B shows  
122 differences among the median values of different groups of patients from a qualitative standpoint.

123 Our data show that strong IgG antibody response was elicited against the identified epitope in non-severe ZIKV  
124 cases and suggests that the antibody response against NS2B is absent in Zika infections with neurological symptoms  
125 (NeuroZIKV) in individuals not previously exposed to DENV. To strengthen this observation, Principal Component  
126 Analysis (PCA) was initially performed to assess the IgG response in the NS2B peptide array dataset. PCA is an  
127 unsupervised technique based on dataset's variance that can be employed as a pattern recognition tool (for further  
128 detail on PCA theory see associated content). The PCA analysis highlighted similarities and differences between  
129 NeuroZIKV and non-severe ZIKV cases, as well as identified the importance of each peptide signal related to these  
130 differences.

131 The scores plot (shown in Figure 2A) shows the relationship among objects analyzed (patients), in such way that  
132 the closer two patients are, more similarities are observed in terms of signal response against the antibody. The 2  
133 first principal components (PCs) obtained from a 2-component PCA model are depicted in Figures 2A and 2B. In  
134 Figure 2A the scores scatter plot is shown for the control group (ZIKV- DENV-), convalescent Zika without Dengue  
135 infection history (ZIKV+, DENV- (group IV)), convalescent Zika with Dengue infection history (ZIKV+, DENV+ (group  
136 V)), convalescent Zika with low antibody titers against Dengue (ZIKV+, DENV low (group VI)) and Zika infection (acute  
137 and convalescent) with neurological symptoms without Dengue infection history (NeuroZIKV, DENV- (group VII))  
138 patients. It is possible to observe that the majority of the samples in the NeuroZIKV, DENV- and the control group  
139 has negative score values for PC1, whereas non-severe ZIKV samples have positive score values of PC1. The loadings  
140 plot, depicted in Figure 2B, shows how each peptide array contributes for the differences observed in the scores  
141 plot. Positive loading values indicates higher intensity responses for patients with positive scores and lower  
142 intensities for patients with negative scores. In PC1, the difference accounts for 37% of total data variance and is  
143 mainly related to the following sequence peptides with positive values of PC1 loading plot (Figure 2B):  
144 GDITWEKDAEVTGNS, GDITWEKDAEITGNS, TWEKDAEVTGNSPRL, TWEKDAEITGNSPRL, KDAEVTGNSPRLDVA and  
145 KDAEITGNSPRLDVA, in which the first four are the peptides described in Figure 1B. All these peptides have in  
146 common the amino acid sequence KDAEV(I)TGNS.

147



148 **Figure 1. Peptide array screening and identification of putative ZIKV NS2B peptides.** (A) Peptide array screening  
 149 showing the selection of 401 peptides covering the ZIKV NS1 and NS2B proteins. The IgM and IgG antibody profiles  
 150 were determined for a control group (n=8) and after incubation of acute and convalescent serum samples from 103  
 151 individuals according to the following groups: i) acute Zika patients without dengue infection history (n=14), ii) acute  
 152 Zika patients with dengue infection history (n=17), iii) acute Zika patients with low antibody titers against dengue  
 153 (n=3), iv) convalescent Zika patients without dengue infection history (n=14), v) convalescent Zika patients with  
 154 dengue infection history (n=17), vi) convalescent Zika patients with low antibody titers against dengue (n=3), vii)  
 155 Zika infections with neurological manifestations (NeuroZIKV) without dengue infection history (n=11), viii)  
 156 NeuroZIKV infections with dengue infection history (n=24). (B) Identification of putative ZIKV NS2B epitopes.  
 157 Immunoreactivity of the ZIKV NS2B peptides was individually assessed against the serum samples groups described  
 158 in Figure 1A. Bars represent the median fluorescence intensity for individual peptides. The four peptides showing  
 159 the highest fluorescence signals in the array are highlighted in red.

160  
161

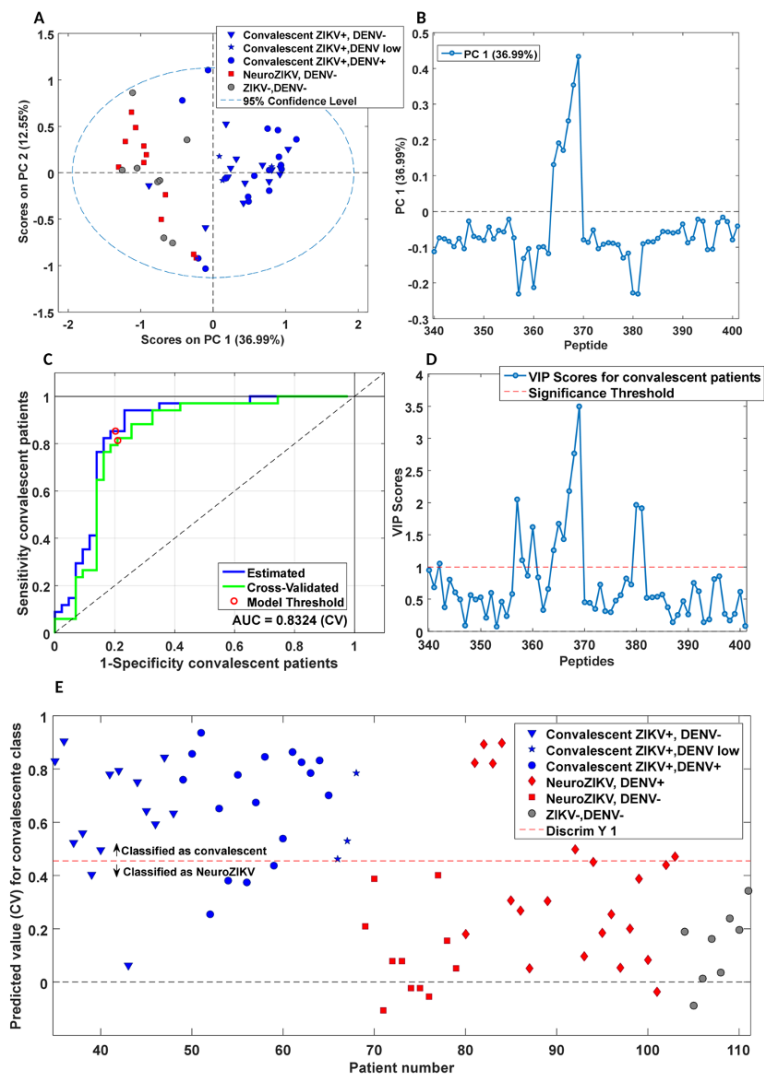
162 According to the loadings and scores plots of PCA model, the abovementioned peptides (positive loading values  
 163 in PC1, Figure 2B) are positively correlated to non-severe ZIKV patients (positive scores in PC1, Figure 2A blue  
 164 symbols) and negatively correlated to NeuroZIKV, DENV- patients (negative scores in PC1, Figure 2A red and grey  
 165 symbols). Therefore, it is possible to state that non-severe ZIKV samples have higher fluorescence signal intensity  
 166 for those peptides when compared to NeuroZIKV, DENV- samples. Moreover, no significant difference among non-  
 167 severe ZIKV patients (groups IV, V and VI) or between NeuroZIKV, DENV- patients and the control group (Figure 2A  
 168 red and grey symbols) were observed. It is important to note that there is no difference between the NeuroZIKV,  
 169 DENV- and the control group, therefore the NeuroZIKV, DENV- signal response is only background noise. Although  
 170 PC2 accounts for 12% of total data variance, its score scatter plot does not show distinctive features.

171 Our results from the PCA model indicate that the previously mentioned peptides are potential candidates to  
 172 discriminate NeuroZIKV patients in the clinical outcomes. To test statistical difference among the groups in a  
 173 multivariate way, the fluorescence signals were employed as classification inputs to build partial least squares  
 174 discriminant analysis (PLS-DA) models. In contrast with PCA, PLS-DA is a supervised technique focused in  
 175 classification and aims to maximize the covariance between the groups of patients and its IgG response profiles (for  
 176 further detail on PLS-DA theory see associated content). In this case, the model is built in a way that a decision

177 function, shown as a threshold on scores plot of PLS-DA (red line Figure 2E), is set to discriminate patients belonging  
 178 to different groups and which peptides are more important in that discrimination.

179 Four different PLS-DA models were built: (i) a global model to test the ability of the epitope of interest to  
 180 discriminate all NeuroZIKV patients (groups VII, VIII) plus the control group from all non-severe ZIKV patients (groups  
 181 IV, V and VI); (ii) a model to discriminate NeuroZIKV, DENV- (group VII) from non-severe ZIKV patients (groups IV, V  
 182 and VI); (iii) a model to discriminate NeuroZIKV (groups VII, VIII) from non-severe ZIKV, DENV- (group IV) patients;  
 183 and (iv) a model to discriminate NeuroZIKV patients according to their DENV background (group VII *versus* VIII).

184 The global model was built to investigate if the differences on IgG signal response for the given epitope will  
 185 suffice to discriminate non-severe (groups IV, V and VI) from NeuroZIKV patients (groups VII and VIII), regardless  
 186 their DENV background. The PLS-DA model shown in Figure 2C provided a value for the area under the curve (AUC)  
 187 of the receiver operating characteristic (ROC) curve equal to 0.83 with sensitivity and specificity values for NeuroZIKV  
 188 class of 79% and 85%, respectively. The variable importance in projection (VIP) scores plot is an important parameter  
 189 to evaluate which peptides are more important to diagnosis prediction (Figure 2D) and must have values above the  
 190 significance threshold. The peptides at data matrix positions 364-367 containing the sequence KDAEV(I)TGNS are  
 191 the ones above the discriminant threshold and, therefore the most important variables to discriminate the two  
 192 groups of patients. This result corroborates that the mentioned peptides are more likely to be an epitope from NS2B  
 193 protein that can be used as a potential biomarker for NeuroZIKV infection.

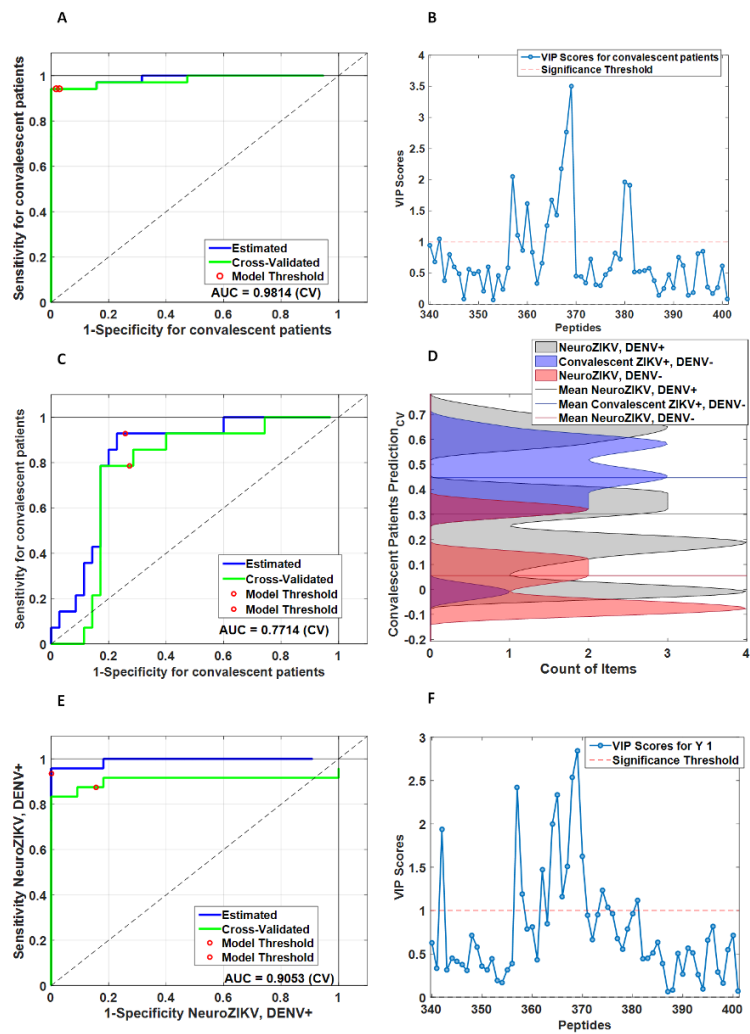


194  
 195 **Figure 2. Summary of the statistical data analysis.** (A) Scores and (B) loadings plots of 2-component PCA model  
 196 including control group, convalescent ZIKV+ (DENV-, DENV+ and DEN low) and NeuroZIKV, DENV- patients; and PLS-  
 197 DA results for the 1-latent variable global model, including (C) ROC curve, (D) VIP scores plot for PLS-DA model and  
 198 (E) predicted values of samples.  
 199

200 Figure 2E shows the predicted scores for each patient, in which those who shows values above 0.45 (defined as  
 201 the threshold by maximizing model sensitivity and specificity) will be classified as convalescent patients. It is worth  
 202 mentioning that thirteen (9 NeuroZIKV DENV+ and 4 non-severe ZIKV+ DENV+) out of fifteen misclassified samples  
 203 belong to patients with previous DENV infection (Figure 2E). This result suggests that antibodies from DENV previous  
 204 infections may cross-react with the identified epitope.

205 To evaluate the influence of NeuroZIKV, DENV+ patients in the classification model performance, a 1-latent  
 206 variable PLS-DA model was built to classify patients as either non-severe or control/NeuroZIKV, DENV-, without  
 207 including NeuroZIKV, DENV+ patients. The obtained ROC curve (Figure 3A) provided an AUC value of approximately  
 208 0.98 at 95% confidence level. This improved classification model showed higher sensitivity and specificity values of  
 209 89% and 94%, respectively, for NeuroZIKV class. Evaluating the VIP scores plot for the classification model (Figure  
 210 3B), it is clear that the peptides containing the KDAEV(I)TGNS sequence appear again as above the significance  
 211 threshold. This result shows that the identified epitope from NS2B protein can be used as a biomarker for the  
 212 identification of ZIKV patients who have developed severe neurological symptoms.

213  
 214  
 215 Interestingly, Figure 1B shows that low IgG antibody response against the identified epitope was observed in  
 216 NeuroZIKV patients that have been previously exposed to DENV. However, this response is apparently higher when  
 217 compared to the control and NeuroZIKV groups without previous DENV infection. This finding suggests that DENV  
 218 infection may elicit IgG response against the NS2B epitope. To further investigate this, an additional PLS-DA model  
 219 was built (Figure 3C and 3D) including only non-severe ZIKV, DENV- and all NeuroZIKV (groups VII, VIII) patients.



220  
 221 **Figure 3. Summary of the additional PLS-DA models.** (A) ROC curve and (B) VIP scores plot of PLS-DA to discriminate  
 222 NeuroZIKV, DENV- from non-severe ZIKV patients (group VII vs groups IV, V and VI); (C) ROC curve and (D)  
 223 Distribution classification histogram of patients on validation step for PLS-DA model to discriminate non-severe ZIKV,  
 224 DENV- (group IV) from NeuroZIKV (groups VII, VIII) patients; (E) ROC curve and (F) VIP scores plot of PLS-DA to  
 225 discriminate NeuroZIKV, DENV- from NeuroZIKV, DENV+ patients (group VII vs group VIII).

226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283

The ROC curve provided an AUC of 0.77 and model sensitivity and specificity for NeuroZIKV class were 71% and 78%, respectively, reflecting a model conflict to find a proper discrimination function. During model optimization step the patients' membership is predicted as a value that range from 0 to 1, with a threshold of 0.33 (defined by maximizing sensitivity and specificity). This means that patients that were predicted with a value higher than 0.33 belong to convalescent class, while values lower than that correspond patients not assigned to convalescent class. It is possible to observe that NeuroZIKV DENV+ patients (grey curve) are predicted along the whole range, while non-severe ZIKV, DENV- patients (blue curve) are systematically above the threshold and NeuroZIKV DENV- patients below the threshold. This result shows a large overlap of the three distributions, particularly NeuroZIKV DENV+ over the other two, which are partially separated from each other. This means that NeuroZIKV patients that were previously exposed to DENV also show significant fluorescence signal and cannot be unequivocally discriminated from non-severe ZIKV+, DENV- patients, in the current dataset. On the other hand, a model built considering only NeuroZIKV patients shows that NeuroZIKV DENV+ are statistically different from NeuroZIKV DENV- cases, and the latter has significantly lower responses against the abovementioned NS2B epitopes (Figure 3E and 3F). Those findings suggest that neurological manifestations of ZIKV infection are associated with reduced antibody response towards the identified NS2B epitope, except when DENV antibodies are present. Corroborating with this finding, microscale thermophoresis (MST) experiments have shown that the serum antibody half-maximal binding parameter (EC50) from the NeuroZIKV samples was remarkably low (EC50 values ca.  $10^{-3}$  Molar), which were comparable to the background noise observed in the negative control samples (Supporting Figures S2A and S2B).

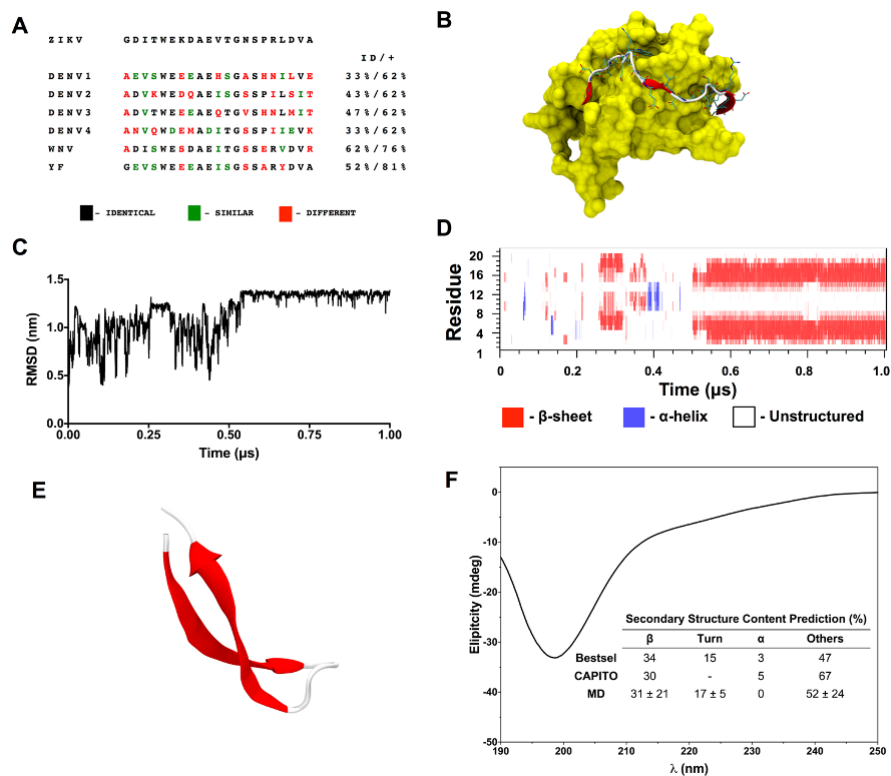
#### **Molecular basis of the interaction between the identified NS2B epitope and ZIKV/DENV antibodies**

Sequence comparison between the identified NS2B epitope and other flaviviruses showed that this region in the NS2B ZIKV protein presents fundamental differences in surface charge (Figure 4A). ZIKV exhibits a positively charged residue (K) in a position where all other flaviviruses exhibit a negatively charged residue (D or E) and an N instead of A, V, or S. In fact, differential charge distribution on the surface of ZIKV and DENV NS1 proteins has previously been documented and suggested as a candidate mechanism to be explored in the query for viral discrimination<sup>29</sup>. Experimental measurements of biomolecular association have shown that hydrogen bonds involving individual charged groups can contribute with up to 4 kcal mol<sup>-1</sup>. This value is considerably lower in uncharged interactions with around 0.5 to 1.5 kcal.mol<sup>-1</sup>. Residues of opposite charge in the ZIKV and DENV epitopes should therefore be central in the discrimination. Interestingly, if charge was assumed to be the main driving force behind discrimination between ZIKV and DENV proteins and related immune response, the expected values for energy association would translate in roughly one order of magnitude binding constant. Microscale thermophoresis (MST) experiments were conducted and have determined the difference in half-maximal binding parameter (EC50) between IgG antibodies in human serum and the identified epitope (Supporting Figures S2C and S2D) was precisely that as determined computationally. However, the measured affinity is more likely to be an apparent affinity (defined here as half-maximum binding parameter given the lack of a 1:1 ratio IgG-peptide) as a total serum polyclonal IgG were included in the experiment. Moreover, despite the punctual charge difference, the overall similarity between ZIKV and DENV NS2B epitope is ca. 62%. Since antibodies against this region appear to be elicited in both DENV and ZIKV non-severe infections, cross-reactivity with non-specific DENV antibodies could be expected, as evidenced in the above PLS-DA models.

It must be kept in mind that antibody responses from a peptide may be suboptimal, as the conformation of peptides may not necessarily reflect of that sequence in the native protein. Peptide flexibility is well known to compromise binding strength and enhance cross-reactivity, which could potentially explain the signals from NeuroZIKV patients previously infected by DENV. In fact, using a less sensitive technique, such as Enzyme-Linked ImmunoSorbent Assay (ELISA), than the peptide array, we have attempted to screen a larger number (N = 42) of serum samples from patients with ZIKV-associated neurological disease using the NS2B peptide to support the finding that IgG antibody response towards the identified epitope is not found in those individuals. However, despite assay optimization, the signal responses were as low as the background noise, which precluded further testing (data not shown for conciseness).

To understand the molecular-level features underlying the interactions between the identified NS2B epitope and the corresponding ZIKV antibodies, we performed sequence and structural analyses of the peptide, the latter by means of molecular dynamics simulations. From a structural point of view, a partial structure of ZIKV NS2B is found in complex with NS3 (Figure 4A), in which the N-terminal region (residues 49–67) of NS2B forms a  $\beta$ -strand packed within the N-terminal region of NS3.<sup>35-38</sup> Although the C-terminal part of NS2B (residues 68–96) also forms a  $\beta$ -hairpin, NMR results have indicated that the C-terminal region shows significant structural diversity, being highly flexible and often disordered.<sup>36, 37</sup> Mishra et al. have found that the lack of the residues K63, D64 drastically decreases the immunogenicity of the peptide.<sup>39</sup> Based on the high flexibility of the C-terminal region and on the finding of Mishra et al., the N-terminal portion of the peptide appears to be a promising site for antibody recognition.

284 Antigenic peptides exist in solution in a variety of conformations, and only conformations that are structurally  
 285 similar to the one in the native protein will bind to the antibody.<sup>40</sup> To characterize the structural dynamics and  
 286 conformational preferences of the 21-mer peptide based on NS2B address (comprising the N-terminal residues 57–  
 287 GDITWEKDAEVTGNSPRLDVA-77 in the NS2B protein), a 1- $\mu$ s molecular dynamics simulation was performed in  
 288 solution. Structural stability was assessed by the positional root-mean-square deviation (RMSD) of all backbone  
 289 atoms, as a function of simulation time (Figure 4B). The native crystal structure (PDB 5H6V<sup>35</sup>) was used as a reference  
 290 structure. The peptide remains disordered and unstructured in the first half of simulation time, as evidenced by a  
 291 large variation on RMSD values and the time-dependent profile of secondary structures (DSSP, Figure 4C). The RMSD  
 292 converges to a value of approx. 1.28 nm after 0.6  $\mu$ s, indicating that the peptide has assumed a preferred  
 293 conformation. The molecule assumes a flexible conformation around a  $\beta$ -hairpin motif, where two  $\beta$ -strands are  
 294 separated by a short loop (Figs 4C and 4D). The antigen-antibody recognition mechanism also requires shape  
 295 complementarity;<sup>41</sup> therefore, highly flexible peptides are unable to efficiently bind to their respective antibodies.<sup>42</sup>  
 296 A stable, close to its native-like conformation would allow the NS2B-based peptide to be recognized by the  
 297 antibodies. The folded structure of the 21-mer peptide in solution, as evidenced by the simulation, shows that only  
 298 parts of the N-terminal residues of the peptide are in a similar conformation as in its native counterpart (Figure S3),  
 299 which can lower the ability of the peptide to bind to the antibody.  
 300



301  
 302 **Figure 4. Sequence and structural analyses of the identified NS2B peptide.** (A) Sequence comparison of the 21-mer  
 303 peptide based on the ZIKV NS2B epitope with the most frequently occurring strains of DENV, WNV, and YFV  
 304 (Residues are color-coded to highlight identity, chemical similarity, and differences. 'ID' stands for identity; a '+'  
 305 indicates similarity degree). (B) Structure representation of ZIKV NS2B (cartoon and sticks) bound to NS3 (yellow  
 306 surface) based on PDB ID 5H6V.<sup>35</sup> (C) Root-mean-square deviation of the NS2B-based peptide as a function of  
 307 simulation time. The same region in PDB ID 5H6V<sup>35</sup> was used as a reference. (D) Secondary structure content of the  
 308 peptide as a function of simulation time. (E) Representative structure of the folded peptide in cartoon ( $\beta$  strands are  
 309 shown in red and unstructured regions in white). (F) Far-UV circular dichroism (CD) spectra of the NS2B peptide in  
 310 acetate buffer collected at 25 °C. Predicted secondary structure content based on experimental data is shown in the  
 311 graph.  
 312

313 To further investigate the secondary structure content of the peptide, we have acquired the experimental  
 314 circular dichroism (CD) spectrum (Figure 4E). The CD graph shows a strong negative band near 200 nm, characteristic  
 315 of disordered polypeptides.<sup>43</sup> Moreover, to obtain a detailed structure information from the CD spectrum, we  
 316 predicted the secondary structure components of the peptide using two different algorithms: BeStSel<sup>44</sup> and  
 317 CAPITO<sup>45</sup>. Both algorithms predicted that the peptide's secondary structure is composed by ca. 30% of  $\beta$ -sheet and



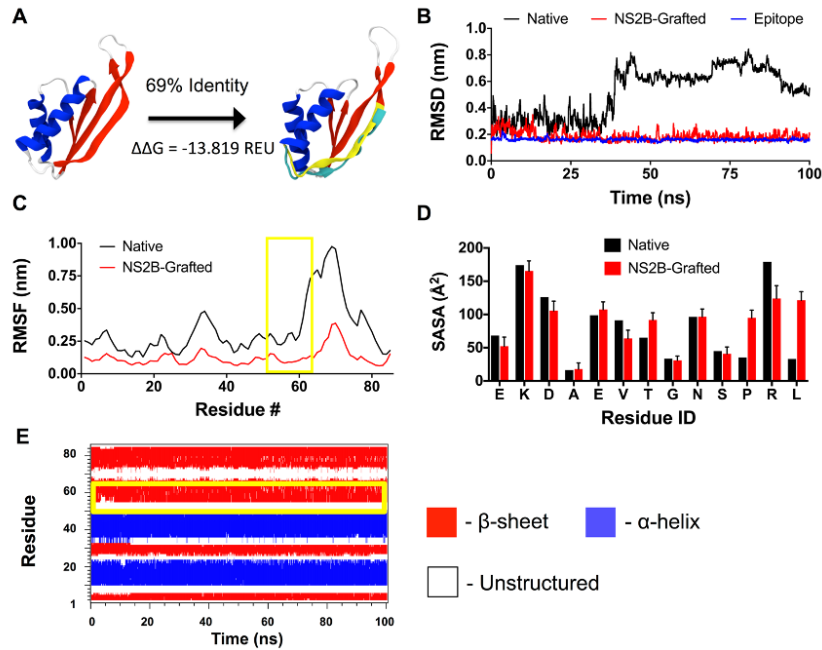
318 ca. 70% of unstructured motifs (Turn + Others) (Figure 4E, inside table). This prediction is in agreement to the MD  
319 data, especially regarding  $\beta$ -sheet content (Figure 4E, inside table). A peptide in solution is characterized by an  
320 ensemble of conformers with different contents of secondary structure and the experimental CD spectrum  
321 represents an average of the contributions from different conformers.<sup>46</sup> Although the MD simulation suggests that  
322 the peptide may adopt a preferred  $\beta$ -hairpin conformation, the variation in its secondary content along the  
323 trajectory shows a significant level of conformational disorder. These results, along with the fact that the peptide  
324 exhibits a CD spectrum typical of disordered proteins, would imply that the unfolded, rather than the folded,  
325 conformers determine the shape of the CD spectra.<sup>43</sup> Together, the findings from MD and CD analyses support that  
326 the peptide in solution adopts a non-native like conformation, leading to a suboptimal affinity by specific ZIKV  
327 antibodies and explaining why the signal responses were not significantly different from background noise in the  
328 ELISA assays and the inability of the peptide array to fully discriminate NeuroZIKV patients with previous DENV  
329 infections from NeuroZIKV/DENV-.

330

### 331 **Engineering an antigen carrying the NS2B epitope and absence of anti-NS2B IgG antibodies in mother-babies paired** 332 **samples**

333 Aiming to overcome the lack of sensibility and specificity of the NS2B peptide due to its high flexibility, we have  
334 performed the computational engineering of a synthetic, soluble and stable protein containing the target epitope.  
335 The new protein was designed to allow the exposure of the epitope in its native conformation, allowing its  
336 recognition by specific antibodies.

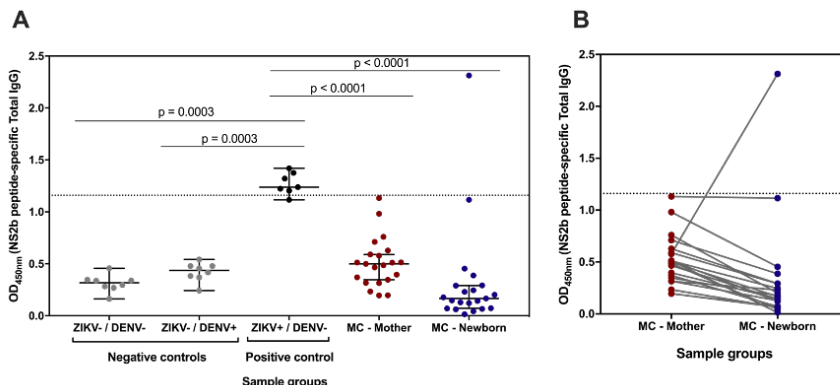
337 Based on Mishra et al.<sup>39</sup> and our findings, we chose to use the residues 62-EKDAEVTGNSPRL-74 of the NS2B  
338 protein as the epitope core. The structural motif of the NS2B epitope was aligned to the structure of 1250 proteins,  
339 yielding one structure with low RMSD and high exposure of the epitope. The scaffold used was the Mja10b protein  
340 (PDB ID: 1NH9),<sup>47</sup> a DNA-binding protein from the hyperthermophilic organism *Methanocaldococcus jannaschii*. In  
341 order to ensure absence of cross-reactivity in the following immunoassays, a BLASTp alignment of the engineered  
342 protein was performed against the entire ZIKV polyproteome showing no identity or even similarity to other regions  
343 than the carried-out epitope (data not shown for conciseness). The epitope was grafted onto the 3<sup>rd</sup>  $\beta$ -sheet (Figure  
344 5A, highlighted in yellow) and after the redesign of the chimeric protein (NS2B-Grafted) the identity relative to native  
345 Mjab10 was 69%. An improvement of  $\Delta\Delta G = -13,819$  REU (Rosetta Energy Units) in the free energy of NS2B-grafted  
346 relative to the native one was also achieved, suggesting that the designed protein is a thermodynamically more  
347 stable structure. Subsequently, the structural stability of the designed protein was evaluated by means of a 100 ns  
348 MD simulation. The native scaffold was also evaluated for comparison purposes. The RMSD for the backbone atoms  
349 of both NS2B-grafted and native scaffold proteins relative to their respective initial structures were calculated over  
350 the entire trajectory (Figure 5B). While the native scaffold showed high RMSD values after 40 ns of simulation time  
351 (Figure 5B, black curve), the NS2B-grafted protein showed values around 0.2 nm throughout the trajectory (Figure  
352 5B, red curve). This result suggests that the NS2B-grafted is structurally more stable, due to the favorable  
353 interactions designed during the modeling stage. The conformational stability of the epitope was also evaluated by  
354 comparing the epitope backbone in the chimera (Figure 5B, blue curve) with its native structure (Figure 4A). The  
355 RMSD of the epitope backbone converges to an average value of 0.17 nm, suggesting a native-like conformation.  
356 This similarity can be seen when the native motif is aligned to the epitope in the NS2B-grafted (Figure 5A, cyan  
357 structure). The mean quadratic fluctuation (RMSF) of the backbone atoms was also evaluated over the entire  
358 trajectory (Figure 5C). For NS2B-grafted protein the RMSF values remained low (under 0.3 nm), including the region  
359 where the epitope was allocated (Figure 5C, highlighted in yellow). The regions with greatest atomic fluctuations  
360 observed correspond to regions of loops and turns, which are expected to be intrinsically disordered. The low  
361 average values of RMSD and RMSF are indicative that the presence of the epitope does not destabilize the scaffold  
362 and the designed NS2B-grafted maintains its native-like structure. The solvent-accessible surface area (SASA) for the  
363 epitope residues was also calculated during the trajectory and compared with the exposure in the native structure  
364 (Figure 5D). Our data suggests that the epitope is exposed, similarly to its the native structure, and available for  
365 antibody recognition. Moreover, this finding includes residues K63 and D64 found by Mishra et al. to be critical for  
366 antigen-antibody binding.<sup>39</sup> In agreement with the RMSD and RMSF data, the DSSP for the NS2B-grafted protein  
367 indicates that all secondary structures were well maintained throughout the simulation (Figure 5E), including the  
368 epitope region (Figure 5E, highlighted in yellow).



369  
 370 **Figure 5. Structural characterization *in silico* of the engineered protein carrying the identified NS2B peptide.** (A)  
 371 Native structure of the scaffold protein (left) and structure of the NS2B-Grafted protein after design/molecular  
 372 dynamics. The percentage indicates the sequence identity between the designed protein and the native scaffold.  
 373  $\Delta\Delta G$  stands for the respective change of free energy of folding after the design protocol in Rosetta units (REU). (B)  
 374 RMSD analysis for the backbone atoms for the Native scaffold (black), for the NS2B-Grafted designed protein (red)  
 375 and comparison of the epitope with its native structure (blue). (C) Flexibility analysis using the RMSF values for the  
 376 backbone atoms of the Native scaffold (black) and NS2B-Grafted (red). The region where the epitope was grafted is  
 377 highlighted in yellow. (D) Solvent accessibility surface (SASA) indicating the exposure per residue of the epitope in  
 378 its native structure (black) and in the designed protein (red). (E) Time-dependent secondary structure map indicating  
 379 the maintenance of structural motifs during the MD simulation.

381 Following the engineering of a protein closely mimicking the natural structure of the NS2B epitope  
 382 (Supplemental Figure S4), presence of IgG towards the NS2B epitope was assessed in a subset of 21 paired samples  
 383 from mothers who delivered babies with ZIKV-associated microcephaly and their children through ELISA. In  
 384 agreement with the computational analysis, the NS2B-grafted protein exhibited enhanced immunoreactivity when  
 385 compared to the NS2B peptide from the peptide array and ELISA (latter not shown for conciseness) and allowed a  
 386 clear distinction between negative and positive control samples (Figure 6A). Corroborating our findings, the anti-  
 387 NS2B IgG antibody profile can be used to distinguish non-severe from severe (NeuroZIKV) infections, while the  
 388 analysis of the paired samples confirmed that this IgG response is absent in the serum from mothers who delivered  
 389 babies with ZIKV-associated microcephaly and their progeny (Figure 6B). On the other hand, the role of the absence  
 390 of such antibodies in the development of neurological outcomes is beyond the scope of this work, and therefore  
 391 remains as an open question.

392



393  
 394

395 **Figure 6. Detection of total IgG antibodies against the NS2B epitope engineered in the NS2B-Grafted protein in**  
396 **paired samples from mothers and babies with ZIKV-associated microcephaly through ELISA.** (A) Scatter plot  
397 representation of NS2B-specific IgG antibodies detection in serum samples from ZIKV naïve (ZIKV-/DENV- (n=8) and  
398 ZIKV-/DENV+ (n=8), negative controls) individuals, ZIKV+/DENV- individuals (n=7, positive controls) and mothers who  
399 delivered babies with ZIKV-associated microcephaly and their progeny (n=21 pairs). The results are shown as mean  
400 values over all measurements and corresponding standard error of the mean (S.E.M.). Dots represent individual  
401 measurements. A dashed line indicates the median of the positive control group for comparison purposes. The p-  
402 values of two-tailed Mann-Whitney test are indicated above each group comparison. (B) Comparison of the  
403 detection of anti-NS2B-IgG in mother-children paired samples. Each sample was tested in duplicate and the mean  
404 values are shown. A dashed line indicates the assay cut-off as calculated from a ROC curve (Figure S5). Gray lines  
405 connect paired samples. MC stands for 'Microcephaly'.

## 406 **Conclusions**

407 We screened well-characterized serum samples from ZIKV-infected individuals in the Northeastern region of Brazil  
408 using high-density peptide arrays displaying the whole proteome of ZIKV. One linear epitope (composed of four  
409 adjacent and partially overlapping peptide sequences) from the ZIKV NS2B protein was identified as a ZIKV-specific  
410 sequence. We found that ZIKV infections elicited high-affinity IgG antibody response against the identified epitope  
411 especially in non-severe ZIKV cases, although previous DENV infections may elicit cross-reactive IgG response against  
412 the identified epitope, when presented as a peptide.

413 Furthermore, our data analysis has shown that a decreased antibody response to this epitope was  
414 associated with severe neurological manifestations after ZIKV infection (NeuroZIKV) compared with uncomplicated  
415 ZIKV infections. Interestingly, unlike what was observed in the uncomplicated ZIKV cases, where the IgG antibody  
416 response did not show any significant difference regardless of DENV background, the NeuroZIKV samples from  
417 patients previously infected with Dengue (NeuroZIKV / DENV+) showed higher IgG antibody responses against NS2B  
418 when compared to NeuroZIKV / DENV- samples. Nevertheless, even though our data indicates that DENV infection  
419 also induces cross-reactive IgG antibodies against the NS2B epitope, the global classification model highlighted the  
420 important contribution of peptides containing the amino acid sequence KDAEV(I)TGNS. Therefore, the IgG antibody  
421 response towards this sequence itself is capable to discriminate between the non-severe ZIKV and NeuroZIKV  
422 infection, regardless of their DENV history, with sensitivity and specificity rates of 79% and 85%, respectively.

423 Our computational and CD analysis for the NS2B peptide in solution indicate that conformational flexibility  
424 as the limiting factor for an unambiguously discrimination of NeuroZIKV patients, as this leads to a lower affinity by  
425 specific ZIKV antibodies as well as cross-reaction with non-specific DENV antibodies. Nevertheless, based on the  
426 native peptide motif, we designed a synthetic protein that mimics the natural structure of the NS2B epitope. Using  
427 this new NS2B-based antigen in an ELISA assay, we have shown that synthetic protein containing the NS2B epitope  
428 can be used as a molecular marker for discrimination between severe and non-severe ZIKV clinical phenotypes.

## 429 **Experimental**

### 430 **Patient population**

431 Sixty-eight (68) cryopreserved serum samples were collected from individuals with acute febrile illness enrolled in a  
432 prospective cohort study from May 2015 to May 2017 established in Recife, Brazil as part of the 'International  
433 Research Consortium on Dengue Risk Assessment, Management, and Surveillance' – IDAMS (ethics committee  
434 protocol number: 15580013.5.1001.5534).<sup>48-50, †</sup> The age of the individuals from whom serum samples were  
435 collected ranged from 9 to 57 years. Sample collection was performed on the first day of recruitment (day 1 – acute  
436 sample), corresponding to the period within the first 72 hours of the febrile phase, and during the convalescent  
437 phase (days 10 – 30 after recruitment – convalescent sample). Furthermore, 35 NeuroZIKV patients enrolled in the  
438 prospective cohort study 'Immunopathology of Dengue infection: viral and host genetic determinants associated to  
439 Dengue pathology' – PRONEX (ethics committee protocol number: 0009.0.095.000-11) were identified around day  
440 1–14 after the onset of neurological symptoms. Nine of the selected samples were collected after this period,  
441 between day 60 and 180 after the onset of neurological symptoms. All samples were classified as convalescent  
442 samples, according to clinical and serological data. Moreover, 23 cryopreserved serum samples were obtained from  
443 subjects enrolled in the clinical cohort 'Prospective study for identification of clinical signs, virological and  
444 immunological factors that are predictive of severe Dengue' – Severe Dengue (ethics committee protocol number:  
445 28309414.9.3001.5201) established in the Department of Virology, Aggeu Magalhães Institute, Oswaldo Cruz  
446 Foundation – FIOCRUZ, Recife, Brazil. Additionally, 21 paired samples from mothers who delivered babies with  
447 microcephaly and their progeny were obtained from subjects enrolled in the cohort study 'Neurological

448 manifestations associated to Zika virus' (ethics committee protocol number: 51106115.8.0000.5190) established  
449 in the Department of Virology, Aggeu Magalhães Institute, Oswaldo Cruz Foundation – FIOCRUZ, Recife, Brazil. We  
450 also included naïve serum samples collected from European donors who had never travelled to tropical areas or  
451 who had visited Brazil for the first time. The characteristics of the patients are shown in Supporting Tables S1 and  
452 S2, while detailed information for each individual NeuroZIKV sample is described on the Supporting Table S3.  
453 Detailed description of the paired samples from mothers who delivered babies with ZIKV-associated microcephaly  
454 and their progeny is provided on Supporting Table S4.

455

#### 456 **Diagnostic testing of participants**

457 All serum samples were tested following the procedures described above. Samples that, according to the molecular,  
458 serological and PRNT tests, were positive only for ZIKV were classified as ZIKV+/DENV-, while samples negative for  
459 all ZIKV tests and positive for at least one DENV serotype were classified as ZIKV-/DENV+. Samples that were negative  
460 for all tested viruses were classified as ZIKV-/DENV-. All patients were either DENV-confirmed or ZIKV-confirmed, or  
461 both. For molecular viral diagnosis, quantitative real-time PCR (qRT-PCR) for DENV and ZIKV was performed  
462 following modified previously reported protocols.<sup>16, 51</sup> As for serology, samples were assayed for anti-DENV IgM and  
463 IgG (Panbio Dengue Capture ELISA) and anti-ZIKV IgM (MAC-ELISA, Centers for Disease Control and Prevention  
464 [CDC])<sup>52</sup> through ELISA. Additionally, all samples were re-tested by indirect DENV ELISA and plaque reduction  
465 neutralization test (PRNT) for ZIKV as well as for each of the four DENV serotypes (DENV1–4) to rule out the  
466 possibility of undetected previous DENV exposure. Within the context of Northeast Brazil, most adults had  
467 previously been exposed to DENV, with seroprevalence rates ranging between 74.3% and 91.1%, also depending on  
468 socio-economic strata.<sup>13</sup> According to the assay results, each sample was classified as: (1) ZIKV+/DENV- if positive  
469 for ZIKV qRT-PCR and negative for DENV qRT-PCR in the acute phase and/or positive for anti-ZIKV IgM with titers >  
470 2 times those for anti-DENV IgM in the convalescent phase in the MAC-ELISA, and negative for anti-DENV IgG in the  
471 acute phase in the Panbio assay; (2) ZIKV-/DENV- if negative for DENV and ZIKV qRT-PCRs in the acute phase,  
472 negative for anti-DENV and anti-ZIKV IgM in the acute phase in MAC-ELISA, and negative for anti-DENV IgG in both  
473 acute and convalescent phases in the Panbio assay<sup>50</sup>. The Panbio Dengue IgG Indirect ELISA was only performed for  
474 the adults and control subgroups. The assessment of virus-specific neutralizing antibodies was performed by PRNT.<sup>50</sup>  
475 Briefly, PRNT assays were performed in Vero cells with locally isolated DENV1–4 and ZIKV strains. PRNT positivity  
476 was calculated based on a 50% reduction in plaque counts (PRNT<sub>50</sub>). Samples were considered positive when  
477 neutralizing antibody levels were ≥1:100 for ZIKV and ≥1:20 against at least one DENV serotype. Specific antibody  
478 titers for ZIKV and DENV were estimated using a four-parameter non-linear regression model. Acute and  
479 convalescent samples were assessed in parallel, and a convalescent/acute titer ratio >4 indicated acute infection.  
480 Important to note that all of the NeuroZIKV samples used in this study has presented anti-ZIKV IgG positive results  
481 through ZIKV IgG ELISA and ZIKV PRNT.

482

#### 483 **Peptide array technology**

484 Whole proteomes of 15 different ZIKV strains (S1 Fig) were retrieved from the National Center for Biotechnology  
485 Information (NCBI) database and translated into overlapping 15-mer peptides (with a peptide-peptide overlap of 12  
486 for DENV and 14 for ZIKV). The accession codes for the ZIKV proteomes from the NCBI were:  
487 ZikaUgandaStrainMR766-NIID, ZikaUgandaStrainMR766, ZikaChinaStrainVE\_Ganxian, ZikaPhilStrainCPC-0740,  
488 ZikaThaiStrainSV0127-14, ZikaChinaStrainZJ03, ZikaHaitiStrain1225/2014, ZikaBrazilStrainSPH2015,  
489 ZikaBrazilStrainNatalRGN, ZikaDomRepStrainPD2, ZikaDomRepStrainPD1, ZikaChinaStrainGD01,  
490 ZikaBrazilStrainZKV2015, ZikaBrazilStrainSSABR1, and ZikaPuertoRicoStrainPRVABC59. The peptide sequences were  
491 aligned, duplicates deleted, and produced as high-density peptide arrays. In addition to that, a c-myc peptide  
492 (EQKLISEEDL) and a peptide derived from poliovirus (KEVPALTAVETGAT, reacts with serum due to comprehensive  
493 vaccination) were included as controls.

494 The arrays were generated in a laser-printer-based approach by PEPperPRINT GmbH (Heidelberg, Germany). Twenty  
495 different amino acid particle types were selectively deposited as spots on a glass slide coated with a poly(ethylene  
496 glycol) methacrylate/poly(methyl methacrylate) graft copolymer functionalized with a βAla-Asp-βAla linker, with  
497 each particle type embedding a different amino acid. When a layer of all twenty different particle types was  
498 completed, the particles were melted at 90 °C, releasing the amino acids containing Fmoc-amino acid  
499 pentafluorophenyl esters that coupled to reactive groups on the functionalized glass slide. Following Merrifield's  
500 principle of orthogonal peptide synthesis, excess material was washed away, and the particle-deposition process  
501 was repeated.<sup>32</sup> In the end, each spot on the glass slide featured a different peptide with a freely chosen sequence.

502

#### 503 **Immunostaining of peptide arrays**

504 Peptide microarrays were placed in incubation trays (PEPperPRINT GmbH; Heidelberg, Germany) and blocked for 30  
505 minutes at room temperature under orbital shaking (140 rotations per minute [rpm]) with blocking buffer MB-070  
506 (Rockland). Human sera were then diluted to 1:1000 in assay buffer (10% Rockland blocking buffer in phosphate-  
507 buffered saline (PBS), 0.05% Tween 20, pH 7.4) and incubated for 16 hours at 4 °C under orbital shaking (140 rpm).  
508 Peptide microarrays were washed three times with PBS supplemented with 0.05% Tween 20 (PBS-T), followed by  
509 incubation with the detection antibodies, goat anti-human IgG Fc specific DyLight 680 conjugated and goat anti-  
510 human IgM antibody DyLight 800 conjugated, diluted to 1:2500 and 1:5000, respectively in assay buffer (PBS-T with  
511 10% Rockland blocking buffer), along with a 1:150 dilution of the control antibody (anti-c-myc antibody,  
512 PEPperPRINT, Germany), for one hour at room temperature under orbital shaking (140 rpm). No specific antibodies  
513 were added to interact with the poliovirus peptide control spots, since poliovirus vaccination in Brazil has been  
514 established across the country. The peptide microarrays were washed three times with PBS-T and rinsed with  
515 deionized water. After drying in a stream of air, fluorescence images were acquired using an Odyssey Imaging System  
516 (LICOR) at 700 and 800 nm, with a resolution of 21  $\mu\text{m}$  and a scanning sensitivity of 7. Image analysis and  
517 quantification were performed with the PepSlide Analyzer software (Sicasys Software GmbH; Heidelberg, Germany).  
518 Background fluorescence was determined as the fluorescence measured when only the secondary antibodies anti-  
519 IgG DyLight 680 (1:2500) and anti-IgM DyLight 800 (1:5000) were added.

#### 520 521 **Statistical analysis**

522 In the present work, multivariate analysis models were built to verify statistical differences among fluorescence  
523 intensities of patient groups. Before building multivariate models, pre-processing techniques were employed to  
524 correct unwanted variation inherent to the measurement. The data set was pre-processed using normalization to  
525 the maximum value of each patient for all variables (peptide fluorescence signals), baseline correction (Automatic  
526 Whittaker Filter with asymmetry of 0.001 and lambda of 100) and mean center. Multivariate data analyses were  
527 carried out with the IgG ZIKV data set in two forms. Initially, the exploratory analysis technique principal component  
528 analysis (PCA)<sup>53</sup> was employed to visualize data variance in a space of reduced dimensionality. Afterwards, a  
529 classification technique based on partial least squares discriminant analysis (PLS-DA)<sup>54</sup> was performed and the  
530 variable importance in projection (VIP scores) plots were employed to evaluate the most important NS2b peptides  
531 to discriminate between patients according their clinical outcomes. Due the small number of samples available and  
532 to include maximum variability to the model, leave-one-out cross-validation method was employed to validate and  
533 determine model complexity based on the true positive rate (or sensitivity (Sn)) and the true negative rate (or  
534 specificity (Sp)). All statistical analyses were performed in PLS\_Toolbox 8.6 (Eigenvector Research Inc., USA) in  
535 MATLAB environment.

#### 536 537 **Engineering an antigen containing the ns2b epitope**

538 The native motif of the NS2B epitope was retrieved from the crystallographic structure deposited in the Protein Data  
539 Bank (PDB ID: 5H6V).<sup>35</sup> The epitope was transplanted, *in silico*, to a carrier protein (scaffold), in order to maximize  
540 its exposure on its native conformation. This procedure was performed using the MotifGraft tool<sup>55</sup> present in the  
541 Rosetta v3.10 software.<sup>56, 57</sup> First, a library of scaffold proteins, composed by 1250 crystallographic structures, was  
542 assembled according to criteria described by Silva *et al.*<sup>55</sup> Then, the NS2B motif was structurally aligned to all  
543 proteins in the library aiming to determine which is the best carrier and region for transplantation. To accept an  
544 alignment, the root mean-square deviation (RMSD) between the backbone of the epitope and the scaffold protein  
545 should be lower or equal to 2.5 Å. The epitope side-chains are transplanted to the scaffold protein if the alignment  
546 criterion is met. After transplantation, the carrier protein is redesigned within a 10 Å radius around the epitope  
547 intending to increase favorable interactions and ensure the folding of the new protein. In order to allow only  
548 mutations that will result in a total stabilizing effect, a position-specific substitution matrix (PSSM) was used,<sup>58</sup>  
549 according to instructions described by Goldenzweig *et al.*<sup>59</sup> PSSM represents the logarithmic probability of looking  
550 any of the 20 possible amino acids in each position. Mutations for the most frequently observed amino acids  
551 generally increase protein stability (the consensus effect), whereas mutations never previously identified in natural  
552 diversity tend to decrease stability.<sup>59-62</sup> During the redesign process, combinatorial sequence optimization is carried  
553 out, with four iterations of sequence design, each one followed by side-chain and backbone minimization. At the  
554 end, the lowest energy structure is obtained according to the REF2015 energy function of the Rosetta.<sup>63</sup>

#### 555 556 **Computational characterization of structural and thermodynamic properties of the identified NS2B peptide and the 557 designed NS2B-grafted antigen**

558 Initial coordinates for the NS2B-based peptide were taken from the crystal structure of the NS2B-NS3 complex (PDB  
559 ID 5H6V) and centered into a cubic box of dimensions 8.0 x 8.0 x 8.0 nm. For the modeled antigen carrying the NS2B

560 epitope, the coordinates were taken from the lowest energy decoy from Rosetta design protocol and centered in a  
561 box of dimensions 12 x 12 x 12 nm. All simulations were carried out by explicitly solvating the biomolecules using  
562 the SPC water model.<sup>64</sup> Periodic boundary conditions were used in the x, y and z directions. Molecular dynamics  
563 simulation was carried out using the NPT ensemble, and all atomistic simulations were performed for 1  $\mu$ s using the  
564 GROMOS 53A6 force field.<sup>65</sup> The LINCS method was used to constrain all bonds.<sup>66</sup> Solute and solvent temperatures  
565 were separately coupled to a thermostat using the velocity rescaling scheme<sup>67</sup> with a reference temperature of  
566 310 K and a relaxation time of 0.2 ps. Pressure was kept constant at 1.013 bar by isotropic coordinate scaling using  
567 the Berendsen barostat<sup>68</sup> with a relaxation time of 0.2 ps and compressibility ( $\kappa_T$ ) of  $4.5 \times 10^{-5}$  bar<sup>-1</sup>. A short-range  
568 cut-off radius of 1.4 nm was used for all non-bonded interactions. Long-range electrostatic interactions were taken  
569 into account using the reaction field method<sup>69</sup> with  $\epsilon=66$  beyond the cut-off radius of 1.4 nm. Counter ions were  
570 added to ensure a neutral system net charge and a salt concentration of 0.150 mM. The systems were first energy-  
571 optimised using 5,000 steps of the steepest descent algorithm. Integration was carried out using an integration time  
572 step of 2 fs based on the leapfrog algorithm.<sup>70</sup> System thermalization was carried out in NVT ensemble for 100 ps,  
573 followed by a pressure equilibration in an NPT ensemble for 500 ps. All equilibration steps were performed with  
574 position constraints for all atoms. All simulations and analyses were performed using the GROMACS 4.6.5 simulation  
575 package.<sup>71</sup> The most representative conformer of each simulated biomolecule was selected by structural clustering  
576 analysis. The clusters were obtained using the GROMOS method<sup>72</sup> implemented in the *g\_cluster* GROMACS tool. A  
577 cut-off of 0.2 nm was used for calculating the root-mean-square deviations (RMSD) of superimposed backbone and  
578  $\beta$ -carbon atoms. For the peptide only the last 550 ns period of the trajectory was used for clustering, whereas for  
579 the modeled antigen containing the NS2B epitope the last 75 ns was used. Visual analyses were carried out using  
580 the Visual Molecular Dynamics (VMD) software version 1.8.7.<sup>73</sup>

581

#### 582 **Circular dichroism**

583 The NS2B peptide was diluted in acetate buffer (0.5 mg/mL) to the concentration of 60  $\mu$ M. Circular dichroism data  
584 were collected on an Olis DSM17 Spectrometer. Scans were recorded in a 0.1 cm path length. Circular dichroism  
585 scans were carried out from 260 to 190 nm with 5 seconds averaging times, 1 nm step size, and 2 nm bandwidth at  
586 25 °C. Spectra were corrected for a buffer blank and baseline molar ellipticity at 260 nm. Scan data were smoothed  
587 by the Stavitsky-Golay method.

588

#### 589 **Detection of anti-NS2B IgG antibodies through ELISA**

590 The DNA sequence coding for the engineered NS2B-grafted protein was optimized for enhanced expression in  
591 prokaryotic systems, cloned into the pET14b (Novagen) expression vector and produced as a 6His-N-terminally  
592 tagged protein. The recombinant protein was purified by automated affinity chromatography followed by a size  
593 exclusion chromatography. High binding, half area 96-well polystyrene plates (Costar; Lowell, MA, USA) were coated  
594 overnight at 4 °C with 1  $\mu$ g/mL of the recombinant protein diluted in 0.2 M carbonate/bicarbonate buffer (Pierce,  
595 IL, USA). Plates were blocked with skimmed milk (Bio-Rad) at 5% (w/v) in PBS-T buffer [1X PBS with 0.05% (v/v)  
596 Tween 20] for 15 minutes at room temperature. Serum samples were diluted (1:50) in assay buffer [5% (w/v)  
597 skimmed milk in PBS-T] and added to plates. All samples were incubated for 2 hours at room temperature. Plates  
598 were washed five times with PBS-T, followed by incubation with the horseradish peroxidase (HRP)-linked antibody  
599 against total IgG (Jackson ImmunoResearch, 1:30 000 dilution) diluted in assay buffer for 1 hour at room  
600 temperature. After a second wash step, the reaction was developed by the addition of tetramethylbenzidine TMB-  
601 KPL substrate (Pierce, IL, USA) for 30 minutes at room temperature, followed by 1N HCl. Optical densities at a  
602 wavelength of 450nm (OD450nm) were read using a microplate spectrophotometer (BioTek; Winooski, VT, USA). A  
603 subset of ZIKV positive and Flavivirus naïve sera were used as positive and negative controls, respectively. The  
604 controls were run in each plate and used to determine the reproducibility of the assay.

#### 605 **Author Contributions**

606 TJ, ETAM, RDL, IFTV and FFL designed the study. FFL and NF performed the peptide microarrays. DFC performed protein  
607 engineering and MD simulations. CSS and RDC performed detailed statistical analysis in the entire data set. IFTV and AFP  
608 produced the engineered protein and performed the CD measurements. IFTV, CMCSA and BHSL performed the MST and  
609 ELISA experiments. MTC, TM and CNLM provided the serum samples. IFTV, RDL and DFC wrote the manuscript with the input  
610 from all authors.

#### 611 **Conflicts of interest**

612 A patent application for the epitope has been filed by TJ, NF, FFL, TM, and ETA. A patent application for the  
613 engineered protein carrying the identified NS2B epitope has been filed by RDL, DFC and IFTV. Otherwise, the authors  
614 have declared that they have no competing interests.

## 615 Acknowledgements

616 This study was supported by grants from the German Research Foundation (DFG) via the Heidelberg Karlsruhe  
617 Research Partnership (HEiKA to TJ and FFL), the European Commission (IDAMS FP7 281803, ZIKAlliance H2020  
618 734548), the German Centre for Infection Research (DZIF), Heidelberg Site, CuraZika Foundation, FACEPE, NUQAAPE,  
619 CAPES, CNPq, INCT-FCx, FAPESP (2019/01255-9) and the German Federal Ministry of Education and Research  
620 (NanoMatFutur grant 13XP5050A). Computer allocation was partly granted by the Brazilian National Scientific  
621 Computing Center (LNCC). The company PEPperPRINT based in Heidelberg supported the design and printing of the  
622 peptide microarrays. The funders had no role in study design, data collection and analysis, decision to publish, or  
623 preparation of the manuscript.

## 624 Notes and references

625 ‡ Cohort Ethics Statement: the human serum samples reported by the authors in this article were provided by the  
626 'International Research Consortium on Dengue Risk Assessment, Management, and Surveillance' – IDAMS,  
627 'Immunopathology of Dengue infection: viral and host genetic determinants associated to Dengue pathology' –  
628 PRONEX, 'Prospective study for identification of clinical signs, virological and immunological factors that are predictive  
629 of severe Dengue' and 'Neurological manifestations associated to Zika virus' established at FIOCRUZ, Recife, Brazil,  
630 after approval by the institutional ethical committee under the protocol numbers 15580013.5.1001.5534,  
631 0009.0.095.000-11, 28309414.9.3001.5201, and 51106115.8.0000.5190, respectively.

- 633 1. N. R. Faria, R. Azevedo, M. U. G. Kraemer, R. Souza, M. S. Cunha, S. C. Hill, J. Theze, M. B. Bonsall, T. A. Bowden,  
634 I. Rissanen, I. M. Rocco, J. S. Nogueira, A. Y. Maeda, F. Vasami, F. L. L. Macedo, A. Suzuki, S. G. Rodrigues, A. C.  
635 R. Cruz, B. T. Nunes, D. B. A. Medeiros, D. S. G. Rodrigues, A. L. N. Queiroz, E. V. P. da Silva, D. F. Henriques, E.  
636 S. T. da Rosa, C. S. de Oliveira, L. C. Martins, H. B. Vasconcelos, L. M. N. Casseb, D. B. Simith, J. P. Messina, L.  
637 Abade, J. Lourenco, L. C. J. Alcantara, M. M. de Lima, M. Giovanetti, S. I. Hay, R. S. de Oliveira, P. D. S. Lemos,  
638 L. F. de Oliveira, C. P. S. de Lima, S. P. da Silva, J. M. de Vasconcelos, L. Franco, J. F. Cardoso, J. Vianez-Junior,  
639 D. Mir, G. Bello, E. Delatorre, K. Khan, M. Creatore, G. E. Coelho, W. K. de Oliveira, R. Tesh, O. G. Pybus, M. R.  
640 T. Nunes and P. F. C. Vasconcelos, *Science*, 2016, **352**, 345.
- 641 2. C. Brito, *Acta Med. Port.*, 2015, **28**, 679.
- 642 3. L. Schuler-Faccini, E. M. Ribeiro, I. M. Feitosa, D. D. Horovitz, D. P. Cavalcanti, A. Pessoa, M. J. Doriqui, J. I. Neri,  
643 J. M. Neto, H. Y. Wanderley, M. Cernach, A. S. El-Husny, M. V. Pone, C. L. Serao, M. T. Sanseverino and F.  
644 Brazilian Medical Genetics Society-Zika Embryopathy Task, *Morb. Mortal. Wkly. Rep.*, 2016, **65**, 59.
- 645 4. D. L. Heymann, A. Hodgson, A. A. Sall, D. O. Freedman, J. E. Staples, F. Althabe, K. Baruah, G. Mahmud, N.  
646 Kandun, P. F. Vasconcelos, S. Bino and K. U. Menon, *Lancet*, 2016, **387**, 719.
- 647 5. WHO, Zika Causality Statement, <https://www.who.int/emergencies/zika-virus/causality/en/>, (accessed  
648 26.03.2019, 2019).
- 649 6. F. Krauer, M. Riesen, L. Reveiz, O. T. Oladapo, R. Martinez-Vega, T. V. Porgo, A. Haeffliger, N. J. Broutet, N. Low  
650 and W. H. O. Z. C. W. Group, *PLoS Med.*, 2017, **14**, e1002203.
- 651 7. Thalia Velho Barreto de Araújo, Prof Ricardo Arraes de Alencar Ximenes, Demócrito de Barros Miranda-Filho,  
652 Wayner Vieira Souza, Ulisses Ramos Montarroyos, Ana Paula Lopes de Melo, Sandra Valongueiro, Maria de  
653 Fátima Pessoa Militão de Albuquerque, Cynthia Braga, Sinval Pinto Brandão Filho, Marli Tenório Cordeiro,  
654 Enrique Vazquez, Danielle di Cavalcanti Souza Cruz, Claudio Maierovitch Pessanha Henriques, Luciana Caroline  
655 Albuquerque Bezerra, Priscila Mayrelle da Silva Castanha, Rafael Dhalia, Ernesto Torres Azevedo Marques-  
656 Júnior, Prof Celina Maria Turchi Martelli, Prof Laura Cunha Rodrigues, o. b. of, i. f. t. M. E. R. Group, t. B. M. o.  
657 Health, t. P. A. H. Organization, I. d. M. I. P. F. Figueira and t. S. H. D. o. Pernambuco, *Lancet Infect. Dis.*, 2017,  
658 **18**, 328.
- 659 8. E. Dirlikov, J. V. Torres, R. B. Martines, S. Reagan-Steiner, G. V. Perez, A. Rivera, C. Major, D. Matos, J. Munoz-  
660 Jordan, W. J. Shieh, S. R. Zaki and T. M. Sharp, *Emerg. Infect. Dis.*, 2018, **24**, 114.
- 661 9. K. Dinnon, E. Gallichotte, E. Fritch, V. Menachery and R. Baric, *PLoS Negl. Trop. Dis.*, 2019, **13**, e0007212.
- 662 10. C. R. Fontes-Garfias, C. Shan, H. Luo, A. E. Muruato, D. B. A. Medeiros, E. Mays, X. Xie, J. Zou, C. M. Roundy,  
663 M. Wakamiya, S. L. Rossi, T. Wang, S. C. Weaver and P. Y. Shi, *Cell Rep.*, 2017, **21**, 1180.
- 664 11. L. Yuan, X. Y. Huang, Z. Y. Liu, F. Zhang, X. L. Zhu, J. Y. Yu, X. Ji, Y. P. Xu, G. Li, C. Li, H. J. Wang, Y. Q. Deng, M.  
665 Wu, M. L. Cheng, Q. Ye, D. Y. Xie, X. F. Li, X. Wang, W. Shi, B. Hu, P. Y. Shi, Z. Xu and C. F. Qin, *Science*, 2017,  
666 **358**, 933.
- 667 12. P. M. S. Castanha, E. J. M. Nascimento, C. Braga, M. T. Cordeiro, O. V. de Carvalho, L. R. de Mendonca, E. A. N.  
668 Azevedo, R. F. O. Franca, R. Dhalia and E. T. A. Marques, *J. Infect. Dis.*, 2017, **215**, 781.
- 669 13. C. Braga, C. F. Luna, C. M. Martelli, W. V. de Souza, M. T. Cordeiro, N. Alexander, F. de Albuquerque Mde, J. C.  
670 Junior and E. T. Marques, *Acta Trop.*, 2010, **113**, 234.

- 671 14. M. R. Duffy, T. H. Chen, W. T. Hancock, A. M. Powers, J. L. Kool, R. S. Lanciotti, M. Pretrick, M. Marfel, S.  
672 Holzbauer, C. Dubray, L. Guillaumot, A. Griggs, M. Bel, A. J. Lambert, J. Laven, O. Kosoy, A. Panella, B. J.  
673 Biggerstaff, M. Fischer and E. B. Hayes, *N. Engl. J. Med.*, 2009, **360**, 2536.
- 674 15. R. Buathong, L. Hermann, B. Thaisomboonsuk, W. Rutvisuttinunt, C. Klungthong, P. Chinnawirotpisan, W.  
675 Manasatienkij, A. Nisalak, S. Fernandez, I. K. Yoon, P. Akrasewi and T. Pliapat, *Am. J. Trop. Med. Hyg.*, 2015, **93**,  
676 380.
- 677 16. R. S. Lanciotti, O. L. Kosoy, J. J. Laven, J. O. Velez, A. J. Lambert, A. J. Johnson, S. M. Stanfield and M. R. Duffy,  
678 *Emerg. Infect. Dis.*, 2008, **14**, 1232.
- 679 17. I. Rodriguez-Barraquer, F. Costa, E. J. M. Nascimento, N. J. Nery, P. M. S. Castanha, G. A. Sacramento, J. Cruz,  
680 M. Carvalho, D. De Olivera, J. E. Hagan, H. Adhikarla, E. A. Wunder, Jr., D. F. Coelho, S. R. Azar, S. L. Rossi, N.  
681 Vasilakis, S. C. Weaver, G. S. Ribeiro, A. Balmaseda, E. Harris, M. L. Nogueira, M. G. Reis, E. T. A. Marques, D.  
682 A. T. Cummings and A. I. Ko, *Science*, 2019, **363**, 607.
- 683 18. W. Dejnirattisai, P. Supasa, W. Wongwiwat, A. Rouvinski, G. Barba-Spaeth, T. Duangchinda, A. Sakuntabhai, V.  
684 M. Cao-Lormeau, P. Malasit, F. A. Rey, J. Mongkolsapaya and G. R. Screaton, *Nat. Immunol.*, 2016, **17**, 1102.
- 685 19. L. Priyamvada, K. M. Quicke, W. H. Hudson, N. Onlamoon, J. Sewatanon, S. Edupuganti, K. Pattanapanyasat,  
686 K. Chokephaibulkit, M. J. Mulligan, P. C. Wilson, R. Ahmed, M. S. Suthar and J. Wrammert, *Proc. Natl. Acad.  
687 Sci. USA*, 2016, **113**, 7852.
- 688 20. J. A. Muller, M. Harms, A. Schubert, B. Mayer, S. Jansen, J. P. Herbeuval, D. Michel, T. Mertens, O. Vapalahti,  
689 J. Schmidt-Chanasit and J. Munch, *Med. Microbiol. Immunol.*, 2017, **206**, 175.
- 690 21. A. Sharma and S. K. Lal, *Front. Microbiol.*, 2017, **8**, 110.
- 691 22. A. Sakudo, A. Viswan, H. Chou, T. Sasaki, K. Ikuta and M. Nagatsu, *Mol. Med. Rep.*, 2016, **14**, 697.
- 692 23. M. Sanchez-Purra, M. Carre-Camps, H. de Puig, I. Bosch, L. Gehrke and K. Hamad-Schifferli, *ACS Infect. Dis.*,  
693 2017, **3**, 767.
- 694 24. D. Focosi, F. Maggi and M. Pistello, *Clin. Infect. Dis.*, 2016, **63**, 227.
- 695 25. A. Chua, I. Prat, C. M. Nuebling, D. Wood and F. Moussy, *PLOS Negl. Trop. Dis.*, 2017, **11**, e0005269.
- 696 26. F. Breitling, A. Nesterov, V. Stadler, T. Felgenhauer and F. R. Bischoff, *Mol. Biosyst.*, 2009, **5**, 224.
- 697 27. C. Katz, L. Levy-Beladev, S. Rotem-Bamberger, T. Rito, S. G. D. Rudiger and A. Friedler, *Chem. Soc. Rev.*, 2011,  
698 **40**, 2131.
- 699 28. F. F. Loeffler, J. Pfeil and K. Heiss, *Methods Mol. Biol.*, 2016, **1403**, 569.
- 700 29. M. C. L. C. Freire, Pol-Fachin, L., Coêlho, D. F., Viana, I. F. T., Magalhães, T., Cordeiro, M. T., Fischer, N., Loeffler,  
701 F. F., Jaenisch, T., Franca, R. F., Marques, E. T. A., Lins, R. D., *ACS Omega*, 2017, **2**, 3913.
- 702 30. L. K. Weber, A. Palermo, J. Kugler, O. Armant, A. Isse, S. Rentschler, T. Jaenisch, J. Hubbuch, S. Dubel, A.  
703 Nesterov-Mueller, F. Breitling and F. F. Loeffler, *J. Immunol. Methods*, 2017, **443**, 45.
- 704 31. L. K. Weber, A. Isse, S. Rentschler, R. E. Kneusel, A. Palermo, J. Hubbuch, A. Nesterov-Mueller, F. Breitling and  
705 F. F. Loeffler, *Eng. Life Sci.*, 2017, **17**, 1078.
- 706 32. V. Stadler, T. Felgenhauer, M. Beyer, S. Fernandez, K. Leibe, S. Guttler, M. Groning, K. König, G. Torralba, M.  
707 Hausmann, V. Lindenstruth, A. Nesterov, I. Block, R. Pipkorn, A. Poustka, F. R. Bischoff and F. Breitling, *Angew.  
708 Chem. Int. Ed.*, 2008, **47**, 7132.
- 709 33. F. F. Loeffler, T. C. Foertsch, R. Popov, D. S. Mattes, M. Schlageter, M. Sedlmayr, B. Ridder, F. X. Dang, C. von  
710 Bojnicic-Kninski, L. K. Weber, A. Fischer, J. Greifenstein, V. Bykovskaya, I. Buliev, F. R. Bischoff, L. Hahn, M. A.  
711 Meier, S. Brase, A. K. Powell, T. S. Balaban, F. Breitling and A. Nesterov-Mueller, *Nat. Commun.*, 2016, **7**, 11844.
- 712 34. P. M. Castanha, M. T. Cordeiro, C. M. Martelli, W. V. Souza, E. T. Marques, Jr. and C. Braga, *Epidemiol. Infect.*,  
713 2013, **141**, 1080.
- 714 35. Y. Li, Z. Zhang, W. W. Phoo, Y. R. Loh, W. Wang, S. Liu, M. W. Chen, A. W. Hung, T. H. Keller, D. Luo and C. Kang,  
715 *Structure*, 2017, **25**, 1242.
- 716 36. A. Roy, L. Lim, S. Srivastava, Y. Lu and J. Song, *PLOS One*, 2017, **12**, e0180632.
- 717 37. Z. Zhang, Y. Li, Y. R. Loh, W. W. Phoo, A. W. Hung, C. Kang and D. Luo, *Science*, 2016, **354**, 1597.
- 718 38. W. W. Phoo, Y. Li, Z. Zhang, M. Y. Lee, Y. R. Loh, Y. B. Tan, E. Y. Ng, J. Lescar, C. Kang and D. Luo, *Nat. Comm.*,  
719 2016, **7**, 13410.
- 720 39. N. Mishra, A. Caciula, A. Price, R. Thakkar, J. Ng, L. V. Chauhan, K. Jain, X. Che, D. A. Espinosa, M. Montoya  
721 Cruz, A. Balmaseda, E. H. Sullivan, J. J. Patel, R. G. Jarman, J. L. Rakeman, C. T. Egan, C. Reusken, M. P. G.  
722 Koopmans, E. Harris, R. Tokarz, T. Briese and W. I. Lipkin, *MBio*, 2018, **9**.
- 723 40. D. H. Sachs, A. N. Schechter, A. Eastlake and C. B. Anfinsen, *Proc. Natl. Acad. Sci. USA*, 1972, **69**, 3790.
- 724 41. A. Datta-Mannan, A. Thangaraju, D. Leung, Y. Tang, D. R. Witcher, J. Lu and V. J. Wroblewski, *MAbs*, 2015, **7**,  
725 483.
- 726 42. I. F. T. Viana, T. A. Soares, L. F. O. Lima, E. T. A. Marques, M. A. Krieger, R. Dhalia and R. D. Lins, *RSC Advances*,  
727 2013, **3**, 11790.
- 728 43. V. Uversky and S. Longhi, *Instrumental Analysis of Intrinsically Disordered Proteins: Assessing Structure and  
729 Conformation*, Wiley, 2011.
- 730 44. A. Micsonai, F. Wien, É. Bulyáki, J. Kun, É. Moussong, Y.-H. Lee, Y. Goto, M. Réfrégiers and J. Kardos, *Nuc. Acids  
731 Res.*, 2018, **46**, W315.
- 732 45. C. Wiedemann, P. Bellstedt and M. Görlach, *Bioinformatics*, 2013, **29**, 1750.
- 733 46. A. Glättli, X. Daura, D. Seebach and W. F. van Gunsteren, *J. Am. Chem. Soc.*, 2002, **124**, 12972.
- 734 47. G. Wang, R. Guo, M. Bartlam, H. Yang, H. Xue, Y. Liu, L. Huang and Z. Rao, *Protein. Sci.*, 2003, **12**, 2815.



- 735 48. T. Jaenisch, Idams, A. Sakuntabhai, Denfree, A. Wilder-Smith and DengueTools, *PLOS Negl. Trop. Dis.*, 2013, **7**,  
736 e2320.
- 737 49. T. Jaenisch, D. T. Tam, N. T. Kieu, T. Van Ngoc, N. T. Nam, N. Van Kinh, S. Yacoub, N. Chanpheaktra, V. Kumar,  
738 L. L. See, J. Sathar, E. P. Sandoval, G. M. Alfaro, I. S. Laksono, Y. Mahendradhata, M. Sarker, F. Ahmed, A.  
739 Caprara, B. S. Benevides, E. T. Marques, T. Magalhaes, P. Brasil, M. Netto, A. Tami, S. E. Bethencourt, M.  
740 Guzman, C. Simmons, N. T. Quyen, L. Merson, N. T. Dung, D. Beck, M. Wirths, M. Wolbers, P. K. Lam, K.  
741 Rosenberger and B. Wills, *BMC Infect. Dis.*, 2016, **16**, 120.
- 742 50. T. Magalhaes, C. Braga, M. T. Cordeiro, A. L. S. Oliveira, P. M. S. Castanha, A. P. R. Maciel, N. M. L. Amancio, P.  
743 N. Gouveia, V. J. Peixoto-da-Silva, Jr., T. F. L. Peixoto, H. Britto, P. V. Lima, A. R. S. Lima, K. D. Rosenberger, T.  
744 Jaenisch and E. T. A. Marques, *PLOS Negl. Trop. Dis.*, 2017, **11**, e0006055.
- 745 51. R. S. Lanciotti, O. L. Kosoy, J. J. Laven, A. J. Panella, J. O. Velez, A. J. Lambert and G. L. Campbell, *Emerg. Infect.*  
746 *Dis.*, 2007, **13**, 764.
- 747 52. M. T. Cordeiro, C. A. Brito, L. J. Pena, P. M. Castanha, L. H. Gil, K. G. Lopes, R. Dhalla, J. A. Meneses, A. C.  
748 Ishigami, L. M. Mello, L. X. Alencar, K. M. Guarines, L. C. Rodrigues and E. T. Marques, *J. Infect. Dis.*, 2016, **214**,  
749 1897.
- 750 53. R. Bro and A. K. Smilde, *Anal. Methods*, 2014, **6**, 2812.
- 751 54. R. G. Brereton and G. R. Lloyd, *J. Chemom.*, 2014, **28**, 213.
- 752 55. D.-A. Silva, B. E. Correia and E. Procko, in *Computational Design of Ligand Binding Proteins*, ed. B. L. Stoddard,  
753 Springer New York, New York, NY, 2016, pp. 285.
- 754 56. A. Leaver-Fay, M. Tyka, S. M. Lewis, O. F. Lange, J. Thompson, R. Jacak, K. W. Kaufman, P. D. Renfrew, C. A.  
755 Smith, W. Sheffler, I. W. Davis, S. Cooper, A. Treuille, D. J. Mandell, F. Richter, Y.-E. A. Ban, S. J. Fleishman, J. E.  
756 Corn, D. E. Kim, S. Lyskov, M. Berrondo, S. Mentzer, Z. Popović, J. J. Havranek, J. Karanicolas, R. Das, J. Meiler,  
757 T. Kortemme, J. J. Gray, B. Kuhlman, D. Baker and P. Bradley, in *Meth. Enzymol.*, eds. M. L. Johnson and L.  
758 Brand, Academic Press, 2011, vol. 487, pp. 545.
- 759 57. S. J. Fleishman, A. Leaver-Fay, J. E. Corn, E.-M. Strauch, S. D. Khare, N. Koga, J. Ashworth, P. Murphy, F. Richter,  
760 G. Lemmon, J. Meiler and D. Baker, *PLOS One*, 2011, **6**, e20161.
- 761 58. S. F. Altschul, E. M. Gertz, R. Agarwala, A. A. Schäffer and Y.-K. Yu, *Nuc. Acid Res.*, 2008, **37**, 815.
- 762 59. A. Goldenzweig, M. Goldsmith, Shannon E. Hill, O. Gertman, P. Laurino, Y. Ashani, O. Dym, T. Unger, S. Albeck,  
763 J. Prilusky, Raquel L. Lieberman, A. Aharoni, I. Silman, Joel L. Sussman, Dan S. Tawfik and Sarel J. Fleishman,  
764 *Mol. Cell*, 2016, **63**, 337.
- 765 60. B. Steipe, B. Schiller, A. Pluckthun and S. Steinbacher, *J. Mol. Biol.*, 1994, **240**, 188.
- 766 61. T. J. Magliery, *Curr. Opin. Struct. Biol.*, 2015, **33**, 161.
- 767 62. L. G. Nivon, S. Bjelic, C. King and D. Baker, *Proteins*, 2014, **82**, 858.
- 768 63. R. F. Alford, A. Leaver-Fay, J. R. Jeliazkov, M. J. O'Meara, F. P. DiMaio, H. Park, M. V. Shapovalov, P. D. Renfrew,  
769 V. K. Mulligan, K. Kappel, J. W. Labonte, M. S. Pacella, R. Bonneau, P. Bradley, R. L. Dunbrack, R. Das, D. Baker,  
770 B. Kuhlman, T. Kortemme and J. J. Gray, *J. Chem. Theory Comput.*, 2017, **13**, 3031.
- 771 64. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans, in *Intermolecular Forces*, ed. B.  
772 Pullman, Springer Netherlands, 1981, vol. 14, ch. 21, pp. 331.
- 773 65. C. Oostenbrink, A. Villa, A. E. Mark and W. F. Van Gunsteren, *J. Compt. Chem.*, 2004, **25**, 1656.
- 774 66. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Compt. Chem.*, 1997, **18**, 1463.
- 775 67. G. Bussi, D. Donadio and M. Parrinello, *The Journal of chemical physics*, 2007, **126**, 014101.
- 776 68. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *The Journal of chemical*  
777 *physics*, 1984, **81**, 3684.
- 778 69. I. G. Tironi, R. Sperb, P. E. Smith and W. F. van Gunsteren, *The Journal of chemical physics*, 1995, **102**, 5451.
- 779 70. R. W. Hockney, in *Methods in Computational Physics*, eds. B. Alder, S. Fernbach and M. Rotenberg, Academic  
780 Press, New York/London, 1970.
- 781 71. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435.
- 782 72. X. Daura, B. Oliva, E. Querol, F. X. Aviles and O. Tapia, *Proteins*, 1996, **25**, 89.
- 783 73. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33.

784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795

796  
797

## SUPPLEMENTAL INFORMATION

### 798 **Identification of a Zika NS2B epitope for which absence of IgG response is** 799 **associated with severe neurological symptoms and the design of a** 800 **biomarker capable of discriminatory diagnostics between severe and non-** 801 **severe clinical phenotypes**

802 Felix F. Loeffler<sup>a,¶</sup>, Isabelle F.T. Viana<sup>b,¶</sup>, Nico Fischer<sup>c</sup>, Danilo F. Coêlho<sup>b,d</sup>, Carolina S. Silva<sup>e</sup>, Antônio F. Purificação  
803 Jr.<sup>b</sup>, Catarina M.C.S. Araújo<sup>b</sup>, Bruno H.S. Leite<sup>b</sup>, Ricardo Durães-Carvalho<sup>f</sup>, Tereza Magalhães<sup>g</sup>, Clarice N.L. Morais<sup>b</sup>,  
804 Marli T. Cordeiro<sup>b</sup>, Roberto D. Lins<sup>2,&</sup>, Ernesto T.A. Marques<sup>b,h,&</sup> and Thomas Jaenisch<sup>c,i,&\*</sup>

805 a. Max Planck Institute of Colloids and Interfaces, Department of Biomolecular Systems, Potsdam, Germany.

806 b. Department of Virology, Aggeu Magalhães Institute, Oswaldo Cruz Foundation, Recife, PE, Brazil.

807 c. Section Clinical Tropical Medicine, Department of Infectious Diseases, Heidelberg University Hospital, Germany.

808 d. Department of Fundamental Chemistry, Federal University of Pernambuco, Recife, PE, Brazil.

809 e. Department of Chemical Engineering, Federal University of Pernambuco, Recife, PE, Brazil.

810 f. Laboratory of Virology, University of Campinas, Campinas, SP, Brazil.

811 g. Arthropod-borne and Infectious Diseases Laboratory (AIDL), Department of Microbiology, Immunology and Pathology,  
812 Colorado State University, Fort Collins, CO, United States of America.

813 h. Department of Infectious Diseases and Microbiology, University of Pittsburgh, Pittsburgh, PA, United States of America

814 i. German Centre for Infection Research (DZIF), Heidelberg Site, Heidelberg, Germany.

815 <sup>¶</sup> These authors contributed equally to this work.

816 <sup>&</sup> These authors also contributed equally to this work.

817 <sup>\*</sup>Corresponding author: thomas.jaenisch@urz.uni-heidelberg.de

818

### 819 Supporting Information – Contents

820

#### 821 **Additional Materials and methods.**

822 **Table S1.** Origin and characterization of the samples used in the peptide array.

823 **Table S2.** Group stratification of the samples used in the MST assays.

824 **Table S3.** Epidemiological, serology and diagnostic information of the severe ZIKV  
825 (NeuroZIKV) samples used in the peptide array.

826 **Table S4.** Laboratorial characterization of the paired serum samples from mothers who  
827 delivered babies with microcephaly and their progeny.

828 **Figure S1.** Sample set used in the peptide array, resulting in the identification of the  
829 NS2B peptide.

830 **Figure S2.** Antibody binding curves to the NS2B peptide by MST.

831 **Figure S3.** Superposition of the folded peptide to its native counterpart as in PD ID  
832 5H6V.

833 **Figure S4.** Flowchart of antigen modeling protocol.

834 **Figure S5.** Sequence alignment among the NS2b-Grafted protein, the native scaffold  
835 and the ZIKV polyprotein.

836 **Figure S6.** Chromatographic purification and SDS-PAGE of the NS2B-Grafted  
837 protein.

838 **Figure S7.** ROC curve analysis of the NS2B-Grafted protein.

## 839 **ADDITIONAL MATERIALS AND METHODS**

### 840 **Antibody affinity determination by microscale thermophoresis (MST)**

841 Twenty micromolar of commercially synthesized NS2B peptide (GenScript,  
842 United States) was labeled using RED fluorescent dye NT-647-NHS Labeling Kit  
843 (NanoTemper Technologies, Germany) according to manufacturer's instructions.  
844 Unreacted dye was removed by buffer-exchange chromatography (EC), and labeled  
845 peptides were eluted in 1X PBS buffer. The total protein content was measured using  
846 a colorimetric assay (Pierce BCA, Protein Assay Kit, United States) in all of the serum  
847 samples used. Binding assays were performed using MST standard coated capillaries  
848 (NanoTemper). MST measurements comparing ZIKV *versus* DENV-infected samples  
849 were conducted with 150 nM of labeled peptide and a series of 16 1:2 serial dilutions  
850 of human serum (where the highest concentration of total IgG was 33.25  $\mu$ M) in 1X  
851 PBS, 0.05% Tween-20 buffer. MST measurements comparing NeuroZIKV samples  
852 with and without previous DENV infection were conducted with 300 nM of labeled  
853 peptide and a series of 16 1:2 serial dilutions of human serum (where the highest  
854 concentration of total IgG was 640 nM) in 1X PBS, 0.05% Tween-20 buffer.  
855 Measurements were performed at 25 °C using the Monolith NT.115 (NanoTemper).  
856 The red excitation LED was set to 40% and laser power to medium. Laser on-time was  
857 set to 30 s, and laser off-time was set to 5 s. The EC50 was derived from three  
858 independent thermophoresis experiments. The quality of each MST run was assessed  
859 by performing a capillary scan before and after each run to check that the fluorescence  
860 between samples remained within 10% variation. The graphs were plotted using Prism  
861 7.0 software.

862

863

864

865

## 866 **Multivariate data analysis**

867 Complex datasets composed by multivariate responses [1], such as the one produced  
868 by the peptide array technology, often require multivariate data analysis techniques to  
869 give a more comprehensive understanding of the data. To get a better assessment of  
870 the relevant information contained in the data, two techniques will be further detailed  
871 here.

872 Principal component analysis (PCA) is an unsupervised learning technique, which  
873 means that it does not require any prior information regarding the class which the  
874 objects analyzed belong. The main goal of PCA analysis is to represent data in a new  
875 space of reduced dimensionality in which the new axes, known as principal  
876 components (PCs), will be built as a linear combination of the original variables (IgG  
877 response profiles) and in the direction of data's maximum variance. In that sense, the  
878 first PC of the model will explain the maximum variance of the data, while the second,  
879 orthogonal to PC1, will explain the maximum variance that is not explained by PC1.  
880 Mathematically speaking, PCA is a decomposition technique in which the  $\mathbf{X}$  matrix  
881 containing measured dataset (patients *versus* IgG response per peptide array) will be  
882 decomposed into two according Equation 1.

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (1)$$

883 The scores matrix  $\mathbf{T}$  contains the new coordinates of samples in the new vector space  
884 of reduced dimensionality, while the loading matrix  $\mathbf{P}^T$  contains the weights of each  
885 original variable in the new PCA model. The residual matrix  $\mathbf{E}$  contains the data  
886 variability that is not included in the model. To a better understanding of PCA and  
887 output interpretation we suggest the following literature.<sup>1-3</sup>

888 In contrast with PCA, that is an unsupervised learning method, partial least  
889 squares discriminant analysis (PLS-DA) is supervised and needs prior information

890 regarding objects to be modeled, in this case the patients' membership. Unlike PCA,  
 891 PLS-DA does not seek for the maximum variance in matrix  $\mathbf{X}$ , but the maximum  
 892 covariance existent between matrix  $\mathbf{X}$  and a  $\mathbf{y}$  vector.<sup>4, 5</sup> The latter consists in a binary  
 893 vector containing objects membership (ones for objects belonging to the class and  
 894 zeros for those that does not).<sup>6</sup> From a mathematical standpoint, a PLS-DA model is  
 895 built according Equations 2 and 3. Although  $\mathbf{T}$  and  $\mathbf{U}$  are score matrices and  $\mathbf{P}$  and  $\mathbf{q}$   
 896 are loadings,  $\mathbf{T}$  and  $\mathbf{P}^T$  are different from the ones in Equation 1, since now those also  
 897 must be good predictors for  $\mathbf{y}$ ;  $\mathbf{E}_x$  and  $\mathbf{e}_y$  are residual matrix and vector, respectively.  
 898 The inner relationship between  $\mathbf{X}$  and  $\mathbf{y}$  is established by Equation 4 and ruled by the  
 899 weight matrix  $\mathbf{W}$ .

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E}_x \quad (2)$$

$$\mathbf{y} = \mathbf{Uq}^T + \mathbf{e}_y \quad (3)$$

$$\mathbf{U} = \mathbf{TW} \quad (4)$$

900 Afterwards, it is possible to obtain the coefficient regression vector  $\mathbf{b}$ , that will be used  
 901 to predict the membership of the objects in a new  $\mathbf{X}_{new}$  matrix as depicted in Equation  
 902 5 and 6. From  $\mathbf{b}$  it is possible to obtain the variable importance in projection (VIP)  
 903 scores that shows which original variables are more important in the discrimination  
 904 model.<sup>7</sup> For a more comprehensive understanding of PLS-DA, the reader is encourage  
 905 to read the referred literature.

$$\mathbf{b} = \mathbf{W}(\mathbf{P}^T\mathbf{W})\mathbf{q}^T \quad (5)$$

$$\hat{\mathbf{y}} = \mathbf{X}_{new}\mathbf{b} \quad (6)$$

906

907

908

## 909 **Antigen modeling procedure**

910           Aiming to design a synthetic, soluble and stable protein carrying the identified  
911 NS2B epitope, we have used the epitope-scaffolding strategy, in which the structural  
912 motif of the epitope is grafted into a larger protein structure. The new protein, in turn,  
913 will expose the epitope in its native conformation, allowing its recognition by specific  
914 antibodies. The stabilization of the epitope's conformation, by incorporating it within a  
915 larger structure, aims to reduce the entropic penalty of when using flexible peptide.<sup>8</sup>

916           The computational engineering of the new antigen was performed using the  
917 MotifGraft<sup>8</sup> tool included in the Rosetta software package.<sup>9</sup> The protocol (Figure S4)  
918 consists of first building a database of scaffold structures. Then, the scaffold library is  
919 computationally searched for possible graft locations, by means of structural  
920 alignments between the scaffold and the epitope motif. If the backbones superimpose  
921 with a RMSD  $\leq 2.5$  Å, and do not produce great steric clashes, the side chains of the  
922 epitope are transplanted to the corresponding positions on the scaffold (side-chain  
923 grafting). However, most natural proteins are only marginally stable,<sup>10</sup> and when  
924 removed from their natural context, many proteins fold incorrectly and aggregate.<sup>11</sup>  
925 Therefore, it is necessary to redesign the grafted protein to ensure correct folding. First,  
926 to keep the epitope in its conformation, we aim to introduce intra-molecular  
927 interactions/contacts that will stabilize it. Thus, all residues within a 10 Å radius around  
928 the epitope are randomly modified, while the epitope itself does not undergo mutation  
929 (only repackaging of its side chains). On the other hand, to keep the folding of the rest  
930 of the scaffold protein (i.e. outside the 10 Å radius), instead of introducing random  
931 mutations, we use a phylogenetic restriction. Based on a database of homologous  
932 sequences to the scaffold protein, a position-specific substitution matrix (PSSM) is  
933 created.<sup>12</sup> Each PSSM element represents the logarithmic probability of observing a  
934 certain amino acid at a specific position in the protein. Thus, during the scaffold

935 redesign, only amino acids with a favorable score are used. The phylogenetic  
936 restriction for the allowed sequence space is based on the well-known fact that, in  
937 general, deleterious mutations are eliminated by natural selection. In other words, the  
938 mutation for the most frequently observed amino acid often increases stability (the  
939 consensus effect).<sup>11</sup> At the end of each mutation step, the structure is minimized and  
940 its energy calculated using a score function.<sup>13</sup> As with free energy ( $\Delta G$ ), a  
941 conformation/sequence with a lower score is more favorable, and the objective is to  
942 obtain the sequence of amino acids with the lower energy to stabilize the protein.<sup>10, 11,</sup>  
943 <sup>13</sup> Thus, the criterion for accepting a mutation is based on thermodynamics free energy  
944 (i.e.  $\Delta\Delta G < 0$ ), in order to increase the stability of both the epitope and the protein as  
945 a whole. The structural dynamics of the final redesigned structure is then characterized  
946 by means of MD.



**Table S1. Origin and characterization of the samples used in the peptide array.**

|                          | DENV previous exposure | Study cohort  | ZIKV RT-PCR (Pos/Neg) | Serological characterization    |                                 |                        |                                  | PRNT              |                 | ZIKV Phenotype    |              |
|--------------------------|------------------------|---------------|-----------------------|---------------------------------|---------------------------------|------------------------|----------------------------------|-------------------|-----------------|-------------------|--------------|
|                          |                        |               |                       | Panbio Dengue IgM Capture ELISA | Panbio Dengue IgG Capture ELISA | CDC Zika IgM MAC-ELISA | Panbio Dengue IgG Indirect ELISA | ZIKV              | DENV            | Acute             | Convalescent |
|                          |                        |               |                       |                                 |                                 |                        |                                  |                   |                 | (< 3 days)        | (> 10 days)  |
|                          |                        |               |                       |                                 |                                 |                        |                                  |                   |                 |                   |              |
| <b>Adults</b>            | Positive               | IDAMS         | 34/0                  | Neg                             | Pos                             | Pos                    | Pos                              | Pos               | 17              | 17                |              |
|                          | Negative               | IDAMS         | 28/0                  | Neg                             | Neg                             | Pos                    | Neg                              | Neg or unspecific | 14              | 14                |              |
|                          | Low positive           | IDAMS         | 6/0                   | Neg                             | Neg                             | Pos                    | Neg                              | Undetermined      | Pos (low titer) | 3                 | 3            |
|                          |                        |               |                       |                                 |                                 |                        |                                  |                   |                 | <b>No disease</b> |              |
| <b>Controls (adults)</b> | Brazilians             | Severe dengue | 0/10                  | Neg                             | Neg                             | Neg                    | Pos                              | Neg               | Pos             | 8                 |              |
|                          | Europeans              | Severe dengue | 0/7                   | Neg                             | Neg                             | Neg                    | Neg                              | Neg               | Neg             | 7                 |              |
|                          |                        |               |                       |                                 |                                 |                        |                                  |                   |                 | <b>NeuroZIKV</b>  |              |
| <b>NeuroZIKV</b>         | Positive               | PRONEX        | 10/17                 | Neg                             | Pos                             | Pos                    | -                                | Pos               | Pos             | 24                |              |
|                          | Negative               | PRONEX        | 7/1                   | Neg                             | Neg                             | Neg                    | -                                | Neg               | Neg             | 11                |              |
|                          | <b>Total</b>           |               |                       |                                 |                                 |                        |                                  |                   |                 | <b>118</b>        |              |

947 **Table S2. Group stratification of the samples used in the MST assays.**

948 <sup>a</sup> Acute Disseminated Encephalomyelitis;

|  | Study cohort  | ZIKV RT-PCR | Neurological diagnosis               | Serological characterization    |                                 |                        |                                  |      |       | N = |
|--|---------------|-------------|--------------------------------------|---------------------------------|---------------------------------|------------------------|----------------------------------|------|-------|-----|
|  |               |             |                                      | Panbio Dengue IgM Capture ELISA | Panbio Dengue IgG Capture ELISA | CDC Zika IgM MAC-ELISA | Panbio Dengue IgG Indirect ELISA | PRNT |       |     |
|  |               |             |                                      |                                 |                                 |                        |                                  | ZIKV | DENV  |     |
| <b>DENV- ZIKV+</b>                       | Severe dengue | Pos         | -                                    | Neg                             | Pos                             | Pos                    | Pos                              | Pos  | Neg   | 3   |
| <b>DENV+ ZIKV-</b>                       | Severe dengue | Neg         | -                                    | Pos                             | Pos                             | Neg                    | -                                | Neg  | -     | 3   |
| <b>NeuroZIKV / DENV-</b>                 | PRONEX        | Pos         | ADEM <sup>a</sup> / GBS <sup>b</sup> | Neg                             | Neg                             | Pos                    | Neg                              | Pos  | Neg   | 2   |
| <b>NeuroZIKV / DENV+</b>                 | PRONEX        | Pos         | GBS                                  | Neg                             | Pos                             | Pos                    | Pos                              | Pos  | Neg   | 2   |
| <b>Negative controls (ZIKV- / DENV-)</b> | PRONEX        | Neg         | -                                    | Neg                             | Neg                             | Neg                    | Neg                              | Neg  | Neg   | 2   |
|  |               |             |                                      |                                 |                                 |                        |                                  |      | Total | 12  |

949 <sup>b</sup> Guillain-Barré Syndrome

950

951

952 **Table S3. Epidemiological, serologic and diagnostic characterization of the severe ZIKV (NeuroZIKV) samples used in the**  
 953 **peptide array.**

| Group                             | Patient Number | Sample ID | Neurological Diagnosis                     | Age | Biological Sex | Days of Symptoms | ZIKV Serology | ZIKV RT-PCR (Serum) | ZIKV RT-PCR (LCR) | DENV Serology |
|-----------------------------------|----------------|-----------|--|-----|----------------|------------------|---------------|---------------------|-------------------|---------------|
| <b>NeuroZIKV / DENV- (N = 8)</b>  | 329            | 416       | Acute disseminated encephalomyelitis       | 2   | F              | 8                | IgM+ IgG-     | Pos                 | Pos               | IgM- IgG-     |
|                                   | 196            | 224       | Guillain-Barré Syndrome                    | 21  | F              | 2                | IgM+          | Neg                 | NT                | IgM- IgG-     |
|                                   | 228            | 268       | Neuritis                                   | 51  | F              | 1                | NT*           | Neg                 | Pos               | IgG-          |
|                                   | 85             | 87        | Guillain-Barré Syndrome                    | 46  | M              | 90               | IgM-          | Pos                 | NT                | IgM- IgG-     |
|                                   | 228            | 235       | Neuritis                                   | 51  | F              | 180              | IgM-          | Pos                 | NT                | IgM- IgG-     |
|                                   | 368            | 402       | Optical Neuritis and demyelinating lesions | 28  | F              | 90               | IgM-          | Pos                 | NT                | IgM- IgG-     |
|                                   | 215            | 219       | Post-arboviral infection dystonia          | 53  | F              | 7                | IgM-          | Pos                 | NT                | IgM- IgG-     |
|                                   | 290            | 315       | Meningoencephalitis                        | 30  | M              | 60               | IgM-          | Pos                 | NT                | IgM- IgG-     |
| <b>NeuroZIKV / DENV+ (N = 27)</b> | 315            | 402       | Encephalitis                               | 18  | M              | 60               | IgM-          | Neg                 | Neg               | IgM- IgG+     |
|                                   | 252            | 303       | Acute disseminated encephalomyelitis       | 56  | F              | 6                | IgM-          | Pos                 | Neg               | IgM- IgG+     |
|                                   | 310            | 397       | Guillain-Barré Syndrome                    | 28  | F              | NA**             | IgM+          | Pos                 | NT                | IgM- IgG+     |
|                                   | 300            | 383       | Guillain-Barré Syndrome                    | 68  | F              | 120              | IgM-          | Pos                 | NT                | IgM- IgG+     |
|                                   | 282            | 358       | Guillain-Barré Syndrome                    | 26  | M              | 14               | IgM+          | Pos                 | Pos               | IgM- IgG+     |
|                                   | 318            | 405       | Guillain-Barré Syndrome                    | 25  | M              | 10               | IgM+          | Pos                 | NT                | IgM- IgG+     |
|                                   | 249            | 297       | Guillain-Barré Syndrome                    | 48  | F              | 4                | IgM+          | Pos                 | Pos               | IgM- IgG+     |
|                                   | 280            | 356       | Guillain-Barré Syndrome                    | 35  | M              | 1                | IgM+          | Pos                 | Pos               | IgM- IgG+     |
|                                   | 189            | 217       | Guillain-Barré Syndrome                    | NA  | F              | 2                | IgM-          | Neg                 | NT                | IgM- IgG+     |
|                                   | 191            | 219       | Guillain-Barré Syndrome                    | NA  | M              | 2                | IgM-          | Pos                 | NT                | IgM- IgG+     |
|                                   | 195            | 223       | Peroneal paralysis and numbness            | 30  | F              | 1                | IgM+          | Neg                 | NT                | IgM- IgG+     |
|                                   | 195            | 294       | Peroneal paralysis and numbness            | 30  | F              | 10               | IgM+          | NT                  | NT                | IgM- IgG+     |
|                                   | 208            | 238       | Guillain-Barré Syndrome                    | 42  | F              | 2                | IgM-          | Pos                 | NT                | IgM- IgG+     |
|                                   | 208            | 293       | Guillain-Barré Syndrome                    | 42  | F              | 9                | IgM+          | NT                  | NT                | IgM- IgG+     |
|                                   | 216            | 315       | Guillain-Barré Syndrome                    | 50  | M              | 4                | IgM+          | Neg                 | NT                | IgM- IgG+     |
|                                   | 290            | 370       | Meningoencephalitis                        | NA  | F              | NA               | IgM+          | Neg                 | NT                | IgG+          |
|                                   | 341            | 429       | Optical Neuritis                           | NA  | F              | NA               | IgM+          | Neg                 | NT                | IgG+          |
|                                   | 296            | 379       | Guillain-Barré Syndrome                    | NA  | F              | NA               | IgM+          | NT                  | Neg               | IgG+          |
|                                   | 331            | 418       | Guillain-Barré Syndrome                    | NA  | F              | NA               | IgM+          | NT                  | NT                | IgG+          |
|                                   | 195            | 294       | Guillain-Barré Syndrome                    | 30  | F              | 1                | IgM+          | Neg                 | NT                | IgM- IgG+     |
|                                   | 274            | 290       | Optical Neuritis                           | 43  | M              | 72               | IgM+          | Neg                 | NT                | IgM- IgG+     |
|                                   | 167            | 171       | Guillain-Barré Syndrome                    | 41  | M              | 60               | IgM-          | Pos                 | NT                | IgM- IgG+     |
|                                   | 322            | 350       | Myasthenia gravis and convulsion           | 56  | M              | 120              | IgM-          | NT                  | NT                | IgM- IgG+     |
|                                   | 343            | 431       | Optical Neuritis                           | NA  | F              | NA               | IgM+          | NT                  | NT                | IgM- IgG+     |
|                                   | 314            | 401       | Guillain-Barré Syndrome                    | 57  | F              | NA               | IgM+          | NT                  | NT                | IgM- IgG+     |
|                                   | 294            | 377       | Guillain-Barré Syndrome                    | 74  | M              | NA               | IgM+          | NT                  | NT                | IgM- IgG+     |
|                                   | 313            | 400       | Guillain-Barré Syndrome                    | 62  | M              | NA               | IgM+          | NT                  | NT                | IgM- IgG+     |

\*NT = Not tested

\*\*NA = Not available

955

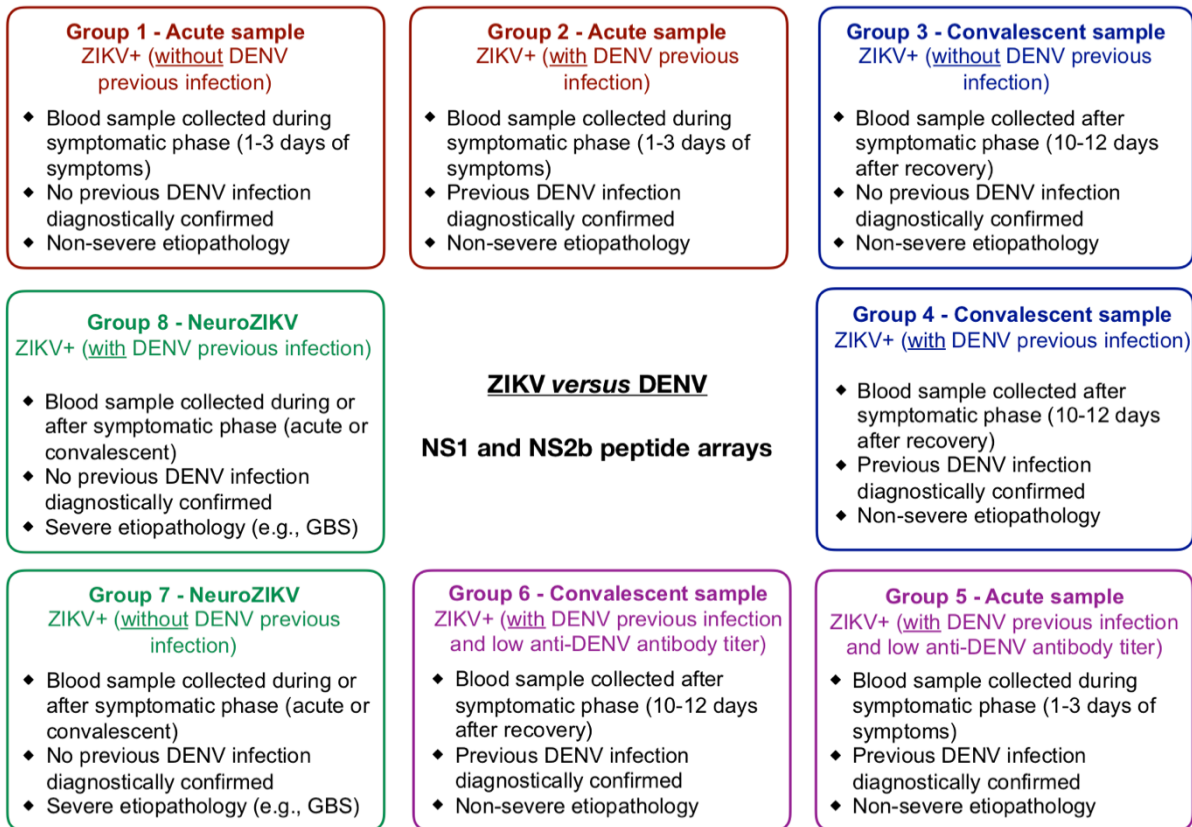
956

957 **Table S4. Laboratorial characterization of the paired serum samples from mothers who delivered babies with microcephaly**  
 958 **and their progeny.**

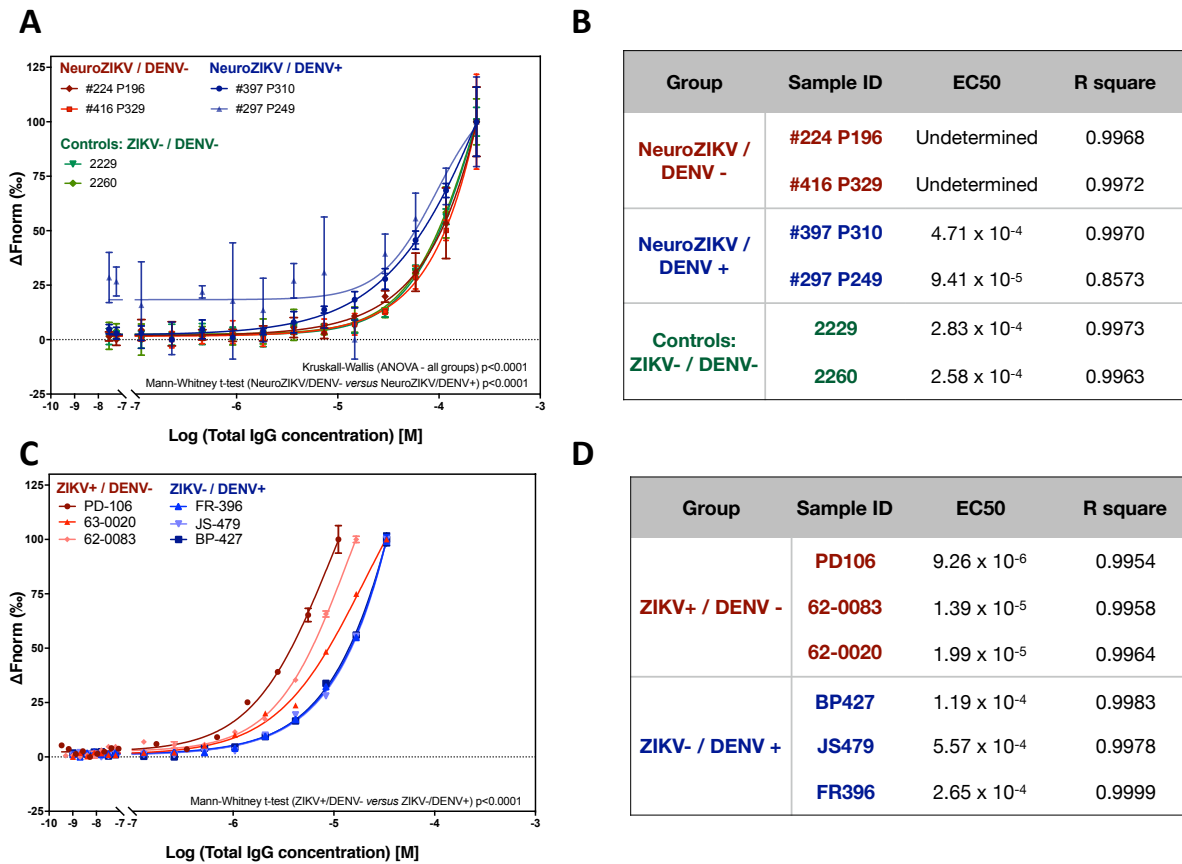
| Paired Samples ID | Biological sex | ZIKV Diagnosis | MC* - Newborn     |                       |                   |                     |            |            |            |            |           |                   | MC - Mother           |            |            |            |            |           |
|-------------------|----------------|----------------|-------------------|-----------------------|-------------------|---------------------|------------|------------|------------|------------|-----------|-------------------|-----------------------|------------|------------|------------|------------|-----------|
|                   |                |                | IgM ZIKV in serum | qRT-PCR ZIKV in serum | IgM ZIKV in CSF** | qRT-PCR ZIKV in CSF | DENV1 PRNT | DENV2 PRNT | DENV3 PRNT | DENV4 PRNT | ZIKV PRNT | IgM ZIKV in serum | qRT-PCR ZIKV in serum | DENV1 PRNT | DENV2 PRNT | DENV3 PRNT | DENV4 PRNT | ZIKV PRNT |
| 01-012-0-1        | M              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Neg        | Pos        | Pos        | Pos       |
| 01-013-0-1        | M              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Neg       | Neg               | Neg                   | Neg        | Neg        | Pos        | Pos        | Neg       |
| 01-014-0-1        | F              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Neg       | Neg               | Neg                   | Neg        | Pos        | Pos        | Neg        |           |
| 01-015-0-1        | M              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Neg       | Neg               | Neg                   | Neg        | Pos        | Pos        | Neg        |           |
| 02-004-0-1        | F              | Yes            | Neg               | Pos                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 02-005-0-1        | F              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 02-006-0-1        | M              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Neg        | Neg        | Neg       | Neg               | Neg                   | Neg        | Neg        | Neg        | Neg        |           |
| 02-008-0-1        | M              | Yes            | Pos               | Neg                   | Pos               | Neg                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 02-018-0-1        | F              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Neg       | Neg               | Neg                   | Neg        | Pos        | Pos        | Neg        |           |
| 03-012-0-1        | M              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 04-003-0-1        | F              | Yes            | Pos               | Neg                   | Pos               | Neg                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 04-004-0-1        | F              | Yes            | Neg               | Neg                   | Pos               | Neg                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 04-006-0-1        | M              | Yes            | Pos               | Neg                   | Pos               | Neg                 | Neg        | Neg        | Pos        | Neg        | Pos       | Neg               | Neg                   | Neg        | Pos        | Neg        | Pos        |           |
| 04-007-0-1        | F              | Yes            | Pos               | Neg                   | Pos               | Neg                 | Neg        | Neg        | Neg        | Neg        | Pos       | Neg               | Neg                   | Neg        | Neg        | Neg        | Pos        |           |
| 04-008-0-1        | F              | Yes            | Pos               | Neg                   | Pos               | Neg                 | Neg        | Neg        | Neg        | Neg        | Pos       | Neg               | Neg                   | Neg        | Neg        | Neg        | Pos        |           |
| 04-019-0-1        | M              | Yes            | Pos               | Neg                   | Pos               | Pos                 | Neg        | Neg        | Neg        | Pos        | Pos       | Neg               | Neg                   | Neg        | Neg        | Pos        | Pos        |           |
| 04-020-0-1        | F              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 06-001-0-1        | M              | Yes            | Pos               | Neg                   | Pos               | Neg                 | Neg        | Neg        | Neg        | Neg        | Pos       | Neg               | Neg                   | Neg        | Neg        | Neg        | Pos        |           |
| 08-001-0-1        | M              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Neg        | Neg        | Pos       | Neg               | Neg                   | Neg        | Neg        | Neg        | Pos        |           |
| 08-003-0-1        | F              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |
| 08-007-0-1        | M              | Yes            | Neg               | Neg                   | Neg               | Pos                 | Neg        | Neg        | Pos        | Pos        | Pos       | Neg               | Neg                   | Neg        | Pos        | Pos        | Pos        |           |

\*MC = Microcephaly  
 \*\*CSF = Cerebrospinal fluid

959

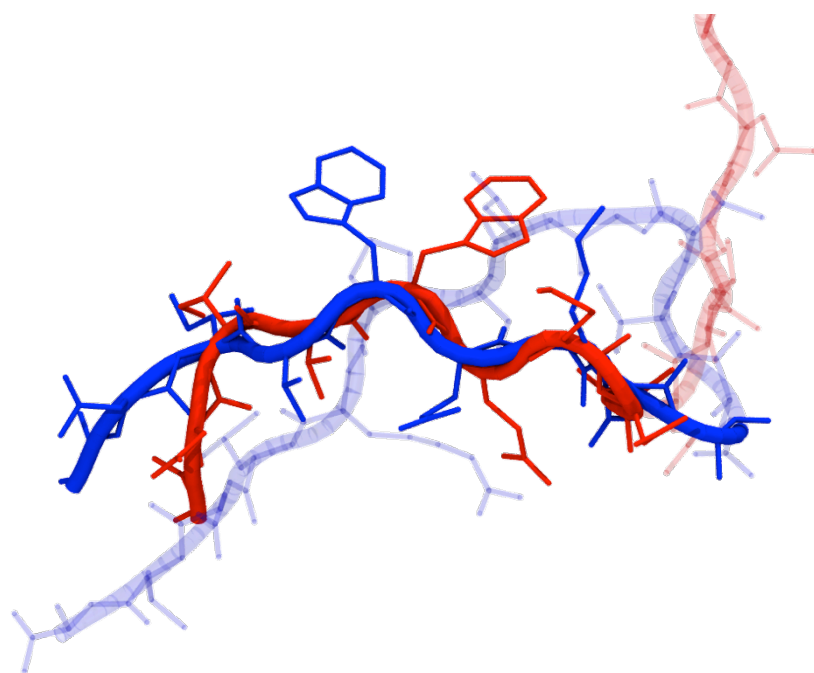


**Figure S1. Sample set used in the peptide array, resulting in the identification of the NS2B peptide.** One hundred and twenty well-characterized serum samples were divided into 8 groups, according to sample stratification molecular and serological tests. Comparison groups included: 1. ZIKV+ acute samples with and without previous DENV infection (coloured in red); 2. ZIKV+ convalescent samples with and without previous DENV infection (coloured in blue); 3. ZIKV+ acute and convalescent samples with low anti-DENV antibody titers (coloured in magenta); 4. NeuroZIKV samples with and without previous DENV infection (coloured in green).



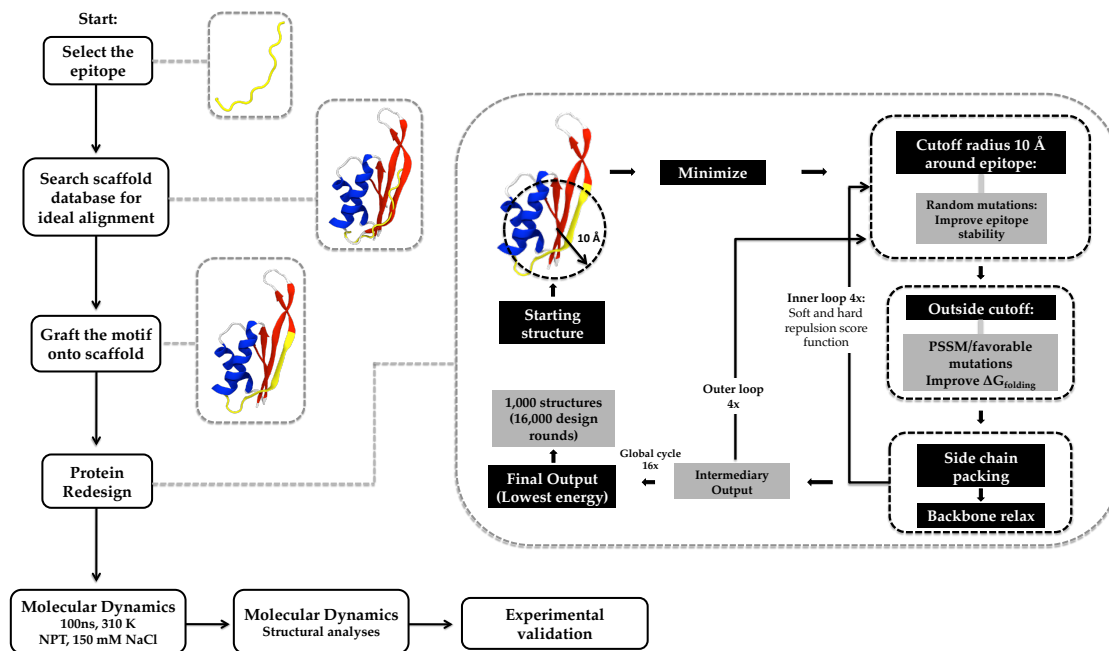
**Figure S2. Antibody binding curves to the NS2B peptide by MST.** (A) Serum binding curves comparison between NeuroZIKV samples with (NeuroZIKV / DENV+) and without previous DENV infection (NeuroZIKV / DENV-). MST data was acquired from serum samples from two subjects exhibiting severe ZIKV infection, with ZIKV RT-PCR and DENV IgG ELISA positive results (NeuroZIKV / DENV+, blue curves) and two subjects exhibiting severe ZIKV infection, with ZIKV RT-PCR positive result and DENV IgG ELISA negative result (NeuroZIKV / DENV-, red curves). Two samples from ZIKV and DENV naïve subjects were included as negative controls (green curves). Plots of the normalised fluorescence ( $\Delta F_{norm}$  (%)) vs. the concentration of total IgG are shown. Error bars represent standard deviations from three individual repeat experiments. The half-maximal binding parameter (EC50) was determined from a non-linear regression analysis. Kruskal-Wallis (ANOVA) test was performed and showed a significant difference among the tested groups ( $p$  value  $< 0.0001$ ). Mann-Whitney  $t$

test between NeuroZIKV/DENV- versus NeuroZIKA/DENV+ samples also revealed a significant difference (p value < 0.0001). (B) Summarizing table indicating the EC50 of anti-NS2B antibodies in the serum samples tested in A. (C) Serum binding curves comparison between ZIKV and DENV positive samples. MST data was acquired from serum samples from three subjects with confirmed single ZIKV infection (ZIKV+/DENV-, red curves) and three subjects with confirmed single DENV infection (ZIKV-/DENV+, blue curves). Plots of the normalised fluorescence ( $\Delta F_{norm}$  (%)) vs. the concentration of total IgG are shown. Error bars represent standard deviations from three individual repeat experiments. The half-maximal binding parameter (EC50) was determined from a non-linear regression analysis. Mann-Whitney t test was performed and showed that a significant difference between the ZIKV+ / DENV- and ZIKV- / DENV+ p-value < 0.0001. (D) Summarizing table indicating the EC50 of anti-NS2B antibodies in the serum samples tested in D.



**Figure S3.** Superposition of the folded peptide (blue) to its native counterpart (red) as in PD ID 5H6V.





**Figure S4.** Flowchart of the antigen modeling protocol using the MotifGraft strategy followed by redesign of the protein to increase epitope conformation and protein folding stability, Molecular Dynamics Analyses and finally experimental validation by means of ELISA assay.

## A

```
NS2b-Grafted      1  N V V Y I G K K P V M N Y V L A V L T Q L N N G T E V V I K A R G K A I S K A V D V A E M V I K R F E K D A S V T G N S P R L D K I T N P D G R E T N V S T I E
Native-Scaffold  1  N V V L I G K K P V M N Y V V A V L T Q L T S N D E V L I K A R G K A I N K A V D V A E M I R N R E I K E I K I K K I E I G T D K V R N P D G R E V N V S T I E

NS2b-Grafted      81  T V L G K
Native-Scaffold   81  T V L A K
```

## B

```
NS2b-Grafted      1  -----N V V Y I G K K F V -----M N Y V L A V L T Q L N N G T E V V I K A R G K A I S K A V D V A E M ---V I K
ZIKV-Capsid       1  K K S G G F R I V N M L K R G V A R V S P F G G L K R L P A G L L L G H G P I R M V L A L A F L R F ---T A I K P S L G L I N R W G S V G K K E A M E I I K

NS2b-Grafted      49  R E E K D A E V T G N S P R L D K I T N P D G R E T N V S T I E I V L G K ---
ZIKV-Capsid       78  K E K K D L A A M L R I I N A R K E K K R R G A D I S V G I V G L L L T A M A
```

## C

```
NS2b-Grafted      1  N V V Y I G K K P V M N Y V L A V L T Q L N N G T E V V I K A R ---G K A I S K A V D V A E M V I K R F E K D A E V T G N S P R L D K I T N P D G R E ---T N
ZIKV-PrM           1  --A E V T R R G S A Y Y M Y --L D R N D A G E A I S F P T T L G M N K C Y I Q I M D L G H M C -----D A T M S Y E C E M L L E G V E P D D V D C W C N

NS2b-Grafted      76  V S T I E I V L G K -----
ZIKV-PrM           71  T T S T W V V Y G T C H H K K G E A R R S R R
```

## D

```
NS2b-Grafted      1  N V V Y I G K K P V M N Y V L A V L T Q L N N G T E V V I K A R G K A I S K A V D V A E M V I K R F E ---K D A E V T G N S P R L D K I T N P D G R E T N V S
ZIKV-M             1  -----V T L P S H S T R K L Q T R S Q T W L E S R E Y T K H I I R V E N W I F R N P G F A L A A A A I A W L L G S S T S Q K V I Y

NS2b-Grafted      78  T I E I V L G K ---
ZIKV-M             63  L V M I L L I A P A Y S
```

## E

```
NS2b-Grafted      1  N V V Y I G -----K K F V M N Y V L A V L T Q L N N G T E V V -----I K A R G
ZIKV-E             1  -I R C L G V S N R D F V E G M S G G T W V D V V L E H G G C V T V M A Q D K E T V D I E L -V T I T V S N M A E V R S Y C Y E A S I S D M A S D S R C P T Q G

NS2b-Grafted      34  K A I S K A V D V A E M V I K R -----E E K D A -----
ZIKV-E             79  E A Y L D K Q S D T Q Y V C K R I L V D R G W G N G C G L E G K G S L V T C A K F A C S K R M T G K S I Q P E N L E Y R I M L S V H G S Q H S G M I V N D T G H

NS2b-Grafted      55  -----E V T G N S P R -----
ZIKV-E            159  E T D E N R A K V E H T E N S P R A E A T L G G F G S L G L D C E P R T G L D F S D L Y Y L T M N N K H W L V H K E W F H D I P L P W H A G A D T G T P H W N N

NS2b-Grafted      63  -----I D K I T -----
ZIKV-E            239  K E A L V E F K D A H A K R Q T V V V L G T Q E G A V H T A L A G A L E A E M D G A K G R L S S G H L K C R L K M D K L R L K G V S Y S L C T A A F T F T K I P

NS2b-Grafted      68  -----N P D G R E T N V S T I -----
ZIKV-E            319  A E T L H G T V T V E V Q Y A G T D G P C K V P A Q M A V D M Q T L I E V G R L I T A N P V I T E S T E N S K M M L E L D P P F G D S Y I V I G V G E K K I T H

NS2b-Grafted      80  -----E I V L G K -----
ZIKV-E            399  H W H R S G S T I G K A F E A T V R G A K R M A V L G D T A W D F G S V G G A L N S L G K G I H Q I F G A A F K S L F G G M S W F S Q L L I G T L L M W L G L N

NS2b-Grafted      -----
ZIKV-E            479  T K N G S I S L M C L A L G G V L I F L S T A V S A
```

## F

NS2b-Grafted 1 -----NVVYIGKKPVMNYVL---  
ZIKV-NS1 1 GCSVDFSKKETRCGTGVFYNDVEAWDRYKYHPDSPRLAAAVKQAWEDGICGISSVSRMENNIMWRSVEGELNATIEEN

NS2b-Grafted 16 -----AVLTQLNNG-----  
ZIKV-NS1 81 GVQLTVVVGSVKNPMWRGPQRLFPVNEIPHCWKAWGKSYFVRAAKTNNFVVDGDTLKECPLKHRANWSFLVEDHGFGV

NS2b-Grafted 25 --TEVVIKAR-----GKAI-----SKAVDVAEMV  
ZIKV-NS1 161 FHTSVWLKVRREDYSLECDPAVIGTAVKGEAVHSDLGWIESEKNDTWRLKRAHLEMTKCEWPKSHTLWTDGIEESDLI

NS2b-Grafted 47 IK-----RFEK-----DAEVTGNSPRLDKIINPDGR-----  
ZIKV-NS1 241 IPKSLAGPLSHHTREGYRTQMKGPHWSEELIRFECPGTKVHVEETCGTRCPSS--RSITASGRVIEEWCCRECTMPPL

NS2b-Grafted 73 -----ETNVSTIEIVLCK  
ZIKV-NS1 320 SFRAKDGCWYGMIEIRPKPEESNLVRSMTAGS

## G

NS2b-Grafted 1 NVVYI-----GKKPVMNYVLAVLTQLNNG-----TEVVIKARGKAIKAVDVAEM-----  
ZIKV-NS2A 1 GVLVLLMQEGLKRRMTTKIIISTSMAVLVAMILCGFSMSDLAKLAILMGATFAEMNTGGDVAHLALIAAFKVRPALLV

NS2b-Grafted 46 -----VIKRFEKDAEVTGNSPRLD-----KITNPDGRETNVSTIEIVLCK  
ZIKV-NS2A 81 SPTFRANWTPRESMLLALASCFLQTAISALEGLMVLINGFALAWLAIRAMVVRRTDNITLAILAALTPL

## H

NS2b-Grafted 1 NVVYIGKKPVMNYVLAVLTQLNNG--TEVVIKARGKAIKAVDVA-----EVIKRI-----EKDAEVTGNSPRLD--  
ZIKV-NS2B 1 SEVLTA-----VGLICALAGGPAKADIEMACPMAAVGLLIVSYVSGKSVDMYIERAGDITWEEKDAEVTGNSPRLDV

NS2b-Grafted 65 -----KITNPDGRETNVSTIEIVL-----GR--  
ZIKV-NS2B 73 ALDESGDFSLVEDDCPPMREIILKVVLMTICGMNPPIAIPFAAGAWYVYVKTGR

## I

NS2b-Grafted 1 -----NVVYI  
ZIKV-NS3 1 SGALWDVPAPKEVKKGETTDGVYRVMTRRLGSTQVGVMQEGVFHTMWHVTKGSALRSGEGRDPYWGDKQLVSYC

NS2b-Grafted 6 GK-----KPMN-----  
ZIKV-NS3 81 CPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGS

NS2b-Grafted 13 -----  
ZIKV-NS3 161 YVSAITQGRREEETPVECFEPSMLKKQLTVLDLHPGAGKTRRVLPPIVREAIKTRLRVILAPTRVVAEMEELRGLP

NS2b-Grafted 13 --YVLAVLTQLNNGTEVVI-----IKARG  
ZIKV-NS3 241 VRYMTTAVNVTHSCTEIVDLMCHATFTRSLLPQIRVFNLYIMDEAHFTDPSIAARGYISTRVEMGEAAAIFMTATPP

NS2b-Grafted 34 -----KAIKAVD-----VAEMVIKRFEKDAEVT--  
ZIKV-NS3 321 GTRDAFPDSNSPIMDTEVEVPERANSSGFDNVTDHSGKTWVWVPSVRNGNEIAACLTKAGKRVIQLSRKTFFETEFQTKH

NS2b-Grafted 58 -----GNSPR-----  
ZIKV-NS3 401 QEWFVVTDISEMGANFKADRVIDSRRCLKPVILDGERVILAGPMPVTHASAAQRGRIGRNENKPGDEYLYGGCAET

NS2b-Grafted 63 -----LKIITNPDG-----RETNVSTI-----  
ZIKV-NS3 481 DEDHAHWLEARMLLNIIYLQDGLIASLYRPEADKVAIEGEFKLRTEQRATFVLMKRGDLPVWLAYQVASAGITYTDRR

NS2b-Grafted 80 -----EIVLCK--  
ZIKV-NS3 561 WCFDGTNTNTIMEDSVPAEVWTRHGEKRVLKPRWMDARVCSHAALKSFKEFFAACKR

## J

NS2b-Grafted 1 NVVYIGKKP-----VMNYVLAVLTQLNNGTEVVIKARGKAIKAVDVAEMV-----IKRFEK  
ZIKV-NS4A 1 VMEALGTLRGHMTERFQEAIDN-LAVLMRAETGSRP-YKAAAAQLPETLETIMLLGLLGTVSLGIFVLMRNKGI GKMGF

NS2b-Grafted 53 DAEVTGNS-----PRLDKITNPDGRETNVSTIEIVLCK-----  
ZIKV-NS4A 79 GMVTIGASAWLMWLSEIEPARIACVLIVVFLLLVLIPEPEKQRSBODNQMAI-IMVAVGLLGLITA

## K

```
NS2b-Grafted      1  NVV-----YIGKK-----FVMNYVL-----AVLTQLNN-----
ZIKV-NS4B        1  NELGWLERTKSDLSHLMGRREEGATIGFSMDIDLRFASAWAIYAALTTFITPAVQHAVTASYQNYSLMAMATQAGVLFGM

NS2b-Grafted     24  -----GTEVVI-----
ZIKV-NS4B       81  GKGMPPFYAWDFGVPLLMIGCYSQLTPLTLIVAIILLVAHYMYLIPGLQAAAAARAAQKRRTAAGIMKNFVVDGIVVTDIDTM

NS2b-Grafted     30  -----KARGKAISKAVDVAEMVIKRF-----EKDAEVIIGNSPRLDKITNED---GREINVESTIEIVLCK-----
ZIKV-NS4B      161  TIDPQVEKMKMGQVLLIAVAVSSAILSRRTAWGWGEAGALITAAATSTLWE-GSPNKYWNSSATSLCNIFRSYLAGASLIY

NS2b-Grafted     ---
ZIKV-NS4B      240  TVTR
```

## L

```
NS2b-Grafted      1  -----
ZIKV-NS5         1  EEDVNLGSGTRAVVSCAEAPNMKIIGNRIERIRSEHAETWFFDENHPYRTWAYHGSYVAPTQGSASSLINGVVRLLSKPW

NS2b-Grafted      1  -----
ZIKV-NS5        81  DVVTGVTGIAMTDTTPYGGQRFVFEKVDTRVPDQEGTRQVMSMVSSWLWKELGKHKRPVCTKEEFINKVRSNAALGAI

NS2b-Grafted      1  -----NVVY--IGKK-----
ZIKV-NS5       161  FEEKEWKTAVEAVNDPRFWALVDKEREHHLRGECQSCVYNNMGRREKKQGEFGKAKGSRAIWYMWLGARFLEFEALGFL

NS2b-Grafted      9  -----PVMNYVLAVLTQLNNG-----TEVVI-----
ZIKV-NS5      241  NEDHWMGRENSGGGVEGLGLQRLGYVLEEMSRIPGRMYADDTAGWDTRISRFDENEAALITNQMEKGRHALALAI IKYT

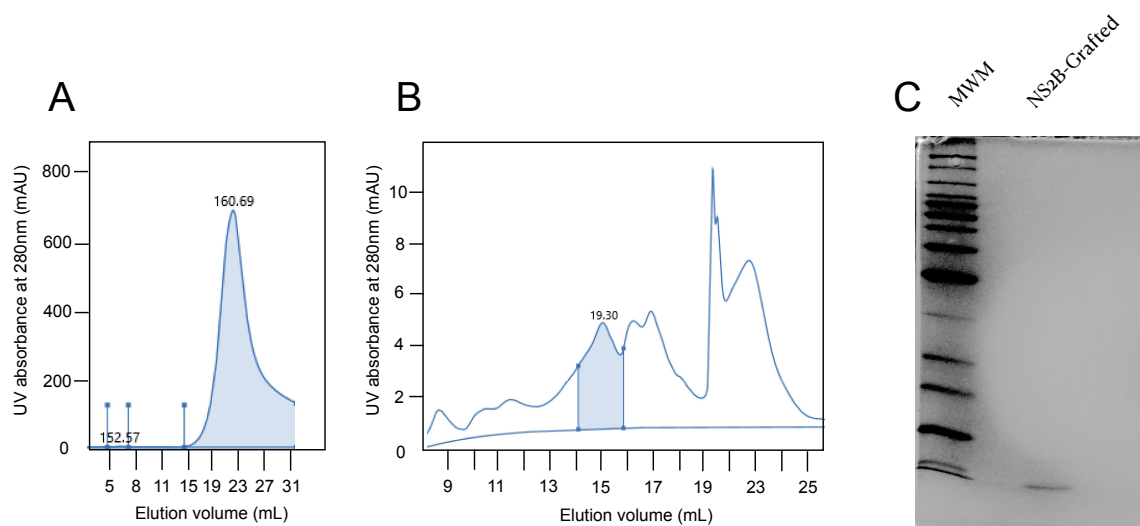
NS2b-Grafted     30  -----KARGKAISKAV-----DVAEM-----VIKRFKDAE-----
ZIKV-NS5      321  YQNKVVVVLRLPAEKGGKTVMDIISRQDQRGSCQVVVYALNFTFNLVVQLIRNMEAEVLEFMQDLWLLRSEKVTNWLQSNQ

NS2b-Grafted     56  -----VIGN-----SPRLDKITNPDGR-----
ZIKV-NS5      401  WDRLLKRMVNSGDDCVVKPIIDDRFAHALRFLNDMGKVRKDTQEWKPKSTGWDNWEVPPFCSHHFNKHLKDGRSIVVPCRHQ

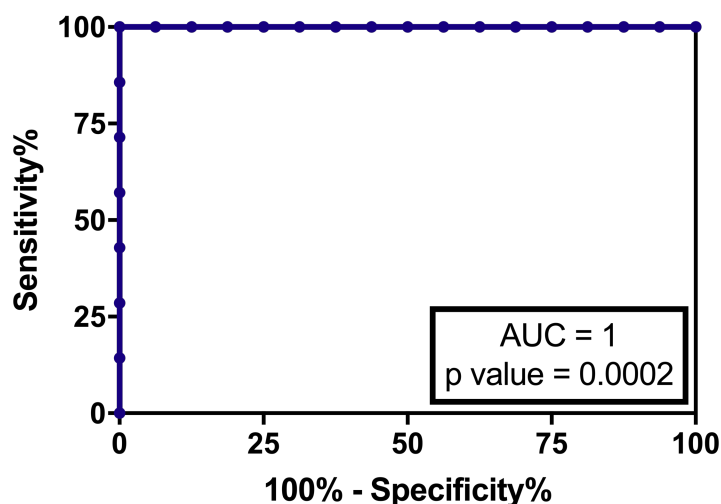
NS2b-Grafted     73  -----
ZIKV-NS5      481  DELIGRARVSPGAGWSIRETACLAKSYAQMWQLLYFHRRDLRLMANAICSSVPVDWVPTGRTTWSIHGKGEWMTTEDMLV

NS2b-Grafted     73  -----ETNVESTIEIVLCK-----
ZIKV-NS5      561  VWNRVWIEENDHMEKTPVTKWTDIPYLGKREDLWCGSLIGHRPRTWAENIKNTVNMVRRITGDEEKYMDYLSLTVQVRY
```

**Figure S5.** Alignment of the NS2b-Grafted protein with (A) the native scaffold; (B) ZIKV Capsid protein; (C) ZIKV PrM protein; (D) ZIKV Membrane protein; (E) ZIKV Envelope protein; (F) ZIKV-NS1 protein; (G) ZIKV-NS2A protein; (H) ZIKV-NS2B protein; (I) ZIKV-NS3 protein; (J) ZIKV-NS4A protein; (K) ZIKV-NS4B protein and (L) ZIKV-NS5 protein. Identical amino acids are black shaded, whereas the epitope is highlighted in yellow.



**Figure S6.** Chromatographic purification and SDS-PAGE of the NS2B-Grafted protein. The NS2B-Grafted protein was submitted to a two-step purification including affinity chromatography (A) followed by size exclusion chromatography (B). UV absorbance at 280 nm (mAU) was monitored over time, and fractions corresponding to the NS2B-Grafted protein (shaded areas in the graphs) were collected. Protein electrophoresis of the NS2B-Grafted protein shows the resultant protein presents a molecular weight of 11.5 kDa, while no aggregates are observed (C). The molecular weight marker (MWM) used was the Benchmark protein ladder (Thermo Fisher Scientific).



**Figure S7.** ROC curve analysis of the NS2B-Grafted protein. The paired results for sensitivity and specificity were plotted as points in a ROC space and the trade-off between these measures for different discrimination cut-offs are graphically represented. The cut-off of 1.16 was used for comparison purposes in the ELISA assay using sera from mothers and babies with ZIKV-associated microcephaly.

## ASSOCIATED CONTENT REFERENCES

1. Bro, R.; Smilde, A. K., Principal component analysis. *Analytical Methods* **2014**, *6* (9), 2812-2831.
2. Ringnér, M., What is principal component analysis? *Nature Biotechnology* **2008**, *26* (3), 303-304.
3. Wold, S.; Esbensen, K.; Geladi, P., Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* **1987**, *2* (1), 37-52.
4. Brereton, R. G.; Lloyd, G. R., Partial least squares discriminant analysis: taking the magic away. *Journal of Chemometrics* **2014**, *28* (4), 213-225.
5. Gromski, P. S.; Muhamadali, H.; Ellis, D. I.; Xu, Y.; Correa, E.; Turner, M. L.; Goodacre, R., A tutorial review: Metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding. *Anal Chim Acta* **2015**, *879*, 10-23.
6. Stähle, L.; Wold, S., Partial least squares analysis with cross-validation for the two-class problem: A Monte Carlo study. *Journal of Chemometrics* **1987**, *1* (3), 185-196.
7. Farrés, M.; Platikanov, S.; Tsakovski, S.; Tauler, R., Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *Journal of Chemometrics* **2015**, *29* (10), 528-536.
8. Silva, D.-A.; Correia, B. E.; Procko, E., Motif-Driven Design of Protein–Protein Interfaces. In *Computational Design of Ligand Binding Proteins*, Stoddard, B. L., Ed. Springer New York: New York, NY, 2016; pp 285-304.
9. Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K. W.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y.-E. A.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popović, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P., Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. In *Methods in Enzymology*, Johnson, M. L.; Brand, L., Eds. Academic Press: 2011; Vol. 487, pp 545-574.
10. Magliery, T. J., Protein stability: computation, sequence statistics, and new experimental methods. *Current opinion in structural biology* **2015**, *33*, 161-168.

11. Goldenzweig, A.; Goldsmith, M.; Hill, Shannon E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; Lieberman, Raquel L.; Aharoni, A.; Silman, I.; Sussman, Joel L.; Tawfik, Dan S.; Fleishman, Sarel J., Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Molecular Cell* **2016**, *63* (2), 337-346.
12. Altschul, S. F.; Gertz, E. M.; Agarwala, R.; Schäffer, A. A.; Yu, Y.-K., PSI-BLAST pseudocounts and the minimum description length principle. *Nucleic Acids Research* **2008**, *37* (3), 815-824.
13. Alford, R. F.; Leaver-Fay, A.; Jeliazkov, J. R.; O'Meara, M. J.; DiMaio, F. P.; Park, H.; Shapovalov, M. V.; Renfrew, P. D.; Mulligan, V. K.; Kappel, K.; Labonte, J. W.; Pacella, M. S.; Bonneau, R.; Bradley, P.; Dunbrack, R. L.; Das, R.; Baker, D.; Kuhlman, B.; Kortemme, T.; Gray, J. J., The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *Journal of Chemical Theory and Computation* **2017**, *13* (6), 3031-3048.