

# Zika virus RNA structure controls its unique neurotropism by bipartite binding to Musashi-1

Received: 17 August 2022

Accepted: 17 February 2023

Published online: 28 February 2023

 Check for updates

Xiang Chen<sup>1,6</sup>, Yan Wang<sup>2,6</sup>, Zhonghe Xu<sup>2,6</sup>, Meng-Li Cheng<sup>1</sup>, Qing-Qing Ma<sup>1</sup>, Rui-Ting Li<sup>1</sup>, Zheng-Jian Wang<sup>1</sup>, Hui Zhao<sup>1</sup>, Xiaobing Zuo<sup>3</sup>, Xiao-Feng Li<sup>1</sup>, Xianyang Fang<sup>2,4</sup> ✉ & Cheng-Feng Qin<sup>1,5</sup> ✉

Human RNA binding protein Musashi-1 (MSI1) plays a critical role in neural progenitor cells (NPCs) by binding to various host RNA transcripts. The canonical MSI1 binding site (MBS), A/GU<sub>(1-3)</sub>AG single-strand motif, is present in many RNA virus genomes, but only Zika virus (ZIKV) genome has been demonstrated to bind MSI1. Herein, we identified the AUAG motif and the AGAA tetraloop in the Xrn1-resistant RNA 2 (xrRNA2) as the canonical and non-canonical MBS, respectively, and both are crucial for ZIKV neurotropism. More importantly, the unique AGNN-type tetraloop is evolutionally conserved, and distinguishes ZIKV from other known viruses with putative MBSs. Integrated structural analysis showed that MSI1 binds to the AUAG motif and AGAA tetraloop of ZIKV in a bipartite fashion. Thus, our results not only identified an unusual viral RNA structure responsible for MSI1 recognition, but also revealed a role for the highly structured xrRNA in controlling viral neurotropism.

Zika virus (ZIKV) infection of pregnant women can cause fetal microcephaly<sup>1</sup>. Interrogation of ZIKV-infected human brain organoids and mouse models demonstrated that ZIKV infection hampers the development of neural progenitor cells (NPCs)<sup>2-5</sup>, which may account for microcephaly in human fetuses or newborn babies. More importantly, ZIKV mainly targets NPCs in the developing brain and replicates with high efficiency in NPCs, whereas mature neurons are less sensitive to ZIKV infection<sup>2,6,7</sup>. This characteristic of ZIKV stands in stark contrast with other known flavivirus members<sup>6</sup>.

RNA-binding protein Musashi homolog 1 (MSI1) is an evolutionarily conserved RNA-binding protein originally discovered in the central nervous system (CNS)<sup>8</sup>. It is highly enriched in undifferentiated neural stem cells (NSCs) and neural progenitor NPCs, but is massively downregulated following the successive progression of neurogenesis<sup>8-10</sup>. MSI1 plays a critical role in maintaining the self-

renewal of NSCs by binding to the 3' untranslated regions (UTRs) of specific target mRNAs<sup>11-14</sup>. Knock out of MSI1 in mice resulted in abnormal brain development and reduced multi-potency of CNS stem cells<sup>15</sup>. MSI1 comprises two RNA recognition motifs (RRM1 and RRM2), connected by a short linker in the N-terminal region followed by an intrinsically disordered C-terminal region. The MSI1 binding site (MBS) has been well characterized as single-stranded RNA motifs containing the consensus sequence A/GU<sub>(1-3)</sub>AG<sup>11,16-19</sup>, and canonical MBSs in various host RNA transcripts encoding proteins with diverse biological functions have been validated<sup>11,20-22</sup>.

The ZIKV genome contains a single open reading frame flanked by 5' and 3' UTRs. The highly structured 3' UTR of 429 nucleotides in length is composed of two xrRNAs (Xrn1-resistant RNA: xrRNA1 and xrRNA2), two dumbbells (DBs, DB1 and DB2), and a 3' stem-loop (SL) subdomain<sup>23</sup>, and is important for viral replication and

<sup>1</sup>Department of Virology, State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, AMMS, Beijing 100071, China. <sup>2</sup>Beijing Advanced Innovation Center for Structural Biology and Beijing Frontier Research Center for Biological Structure, School of Life Sciences, Tsinghua University, Beijing 100084, China. <sup>3</sup>X-ray Science Division, Argonne National Laboratory, Lemont, IL 60439, USA. <sup>4</sup>Key Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. <sup>5</sup>Research Unit of Discovery and Tracing of Natural Focus Diseases, Chinese Academy of Medical Sciences, Beijing 100071, China. <sup>6</sup>These authors contributed equally: Xiang Chen, Yan Wang, Zhonghe Xu.

✉ e-mail: [fangxy@ibp.ac.cn](mailto:fangxy@ibp.ac.cn); [qincf@bmi.ac.cn](mailto:qincf@bmi.ac.cn)

pathogenesis<sup>24,25</sup>. Recently, Chavali et al.<sup>26</sup> demonstrated that MSII directly binds to the 3' UTR of the ZIKV genome and promotes viral replication in cells expressing MSII. Sequence analysis and computational prediction<sup>27–29</sup> have predicated multiple putative MBSs (pMBSs) in the 3' UTR of ZIKV, although none of these pMBSs have been validated. More importantly, these pMBSs are distributed in a large number of viruses across different families with different properties, especially mosquito-borne flaviviruses (MBFV)<sup>26,29</sup>. However, ZIKV is the only known pathogen linked to human neurodevelopmental disorders. This unusual phenotype questions the canonical rule for MSII binding, and the authentic viral RNA targets of MSII demands urgent investigation.

In this study, using combined technology platforms, we identified two MBSs in the xrRNA2 of the ZIKV 3' UTR, and clarified their crucial roles during ZIKV replication in NPCs. Structural analysis using integrated methods demonstrated a unique bipartite binding mode for both xrRNA2 and MSII. Our study not only characterized the evolutionally conserved RNA elements in the ZIKV genome that cooperatively regulate viral replication in diverse cells, but also provides a novel example showing how a virus adopts its RNA primary sequence and tertiary structure to hijack a host protein to facilitate its specific tropism and life cycle.

## Results

### The AUAG motif in ZIKV xrRNA2 is the authentic RNA target of MSII

Human MSII, constituted by RRM1 and RRM2, directly binds to the ZIKV 3' UTR<sup>26</sup>. To characterize the role of each RRM of MSII in RNA-protein interactions, isothermal titration calorimetry (ITC) assays were performed using the tandem RRM12 (20–191) and the individual RRM1 (20–95) and RRM2 (108–191) sequences of MSII and the ZIKV 3' UTR (Fig. S1a, b). As expected, the tandem RRM12 shows obvious binding to the ZIKV 3' UTR with a binding affinity of 2.45  $\mu$ M, while the individual RRM1 or RRM2 subdomains failed to bind with the ZIKV 3' UTR (Fig. S1c). This result suggests that both the RRM1 and RRM2 subdomains of MSII are required for cooperative binding to the ZIKV 3' UTR.

Based on the canonical A/GU<sub>(1–3)</sub>AG rule, three pMBSs (pMBS1–3) have been in silico mapped to the xrRNA2, DB12, and 3' SL subdomains of the ZIKV 3' UTR, respectively (Fig. 1a). In order to clarify the role of each pMBS, the individual subdomains of xrRNA2, DB12 and the 3' SL harboring pMBS1, pMBS2, and pMBS3, respectively, were constructed, and their interactions with MSII RRM12 were measured by ITC. Remarkably, only xrRNA2 showed high binding affinity (1.76  $\mu$ M) to MSII; whilst the binding affinity of the 3' SL/DB12 and MSII was weak (Fig. 1b). Furthermore, three mutants of the ZIKV 3' UTR, named as pMBS1m, pMBS2m, and pMBS3m, respectively, were constructed, in which the respective pMBSs were disrupted without altering the 3' UTR secondary structure (Fig. 1c). Small angle X-ray scattering (SAXS) data revealed that the mutations didn't alter the overall structure of the ZIKV 3' UTR (Fig. S1d–f and Table S1). ITC assays showed that only the mutations contained in pMBS1m significantly weakened MSII binding, whilst the mutated pMBS2m and pMBS3m structures had similar binding affinity as the wild type (WT) 3' UTR (Fig. 1d and Table S2). RNA pull-down assays further confirmed that the mutations in pMBS1m significantly decreased the binding to MSII which was not the case for either pMBS2m or pMBS3m (Fig. 1e). As a control, we also detected whether these three mutations affect the interaction of 3' UTR with Fragile X Mental Retardation Protein (FMRP), which was previously reported to interact with ZIKV 3' UTR<sup>30</sup>. As expected, there was no observed difference in FMRP binding between WT 3'UTR RNA and the three mutant RNAs (Fig. 1e), suggesting the pMBS1m mutation harbors a specific effect on MSII binding. Together, these results demonstrate that pMBS1, the AUAG motif, in xrRNA2 is crucial for MSII binding.

To further determine the function of pMBS1 during ZIKV infection, the corresponding pMBS1m mutation was engineered into the infectious clone of ZIKV strain FSS13025<sup>31</sup> and the corresponding mutant virus (pMBS1m) rescued (Fig. S2a). As expected, ZIKV WT and pMBS1m exhibited similar plaque morphologies (Fig. S2b) and replication kinetics in BHK-21 cells, suggesting that the pMBS1m mutation does not affect viral replication in the absence of MSII (Fig. 1f). In addition, the northern blot result showed that the pMBS1m mutation has no influence on the production of sfRNAs (Fig. S2c). More importantly, in BHK-21 cells that transiently expressed MSII, RNA immunoprecipitation (RNA-IP) assays showed a remarkable reduction in the binding of MSII to viral RNA following ZIKV pMBS1m infection, compared to the WT virus (Fig. 1g). In contrast, the control protein FMRP showed similar binding ability to WT ZIKV and pMBS1m ZIKV RNA (Fig. S2d). Consistently, the co-localization of MSII and viral RNA was significantly reduced in ZIKV pMBS1m infected cells (Fig. 1h). Taken together, our results indicated that the AUAG motif in the ZIKV xrRNA2 is the authentic RNA target of MSII.

### The pMBS1 promotes ZIKV replication in an MSII-dependent manner

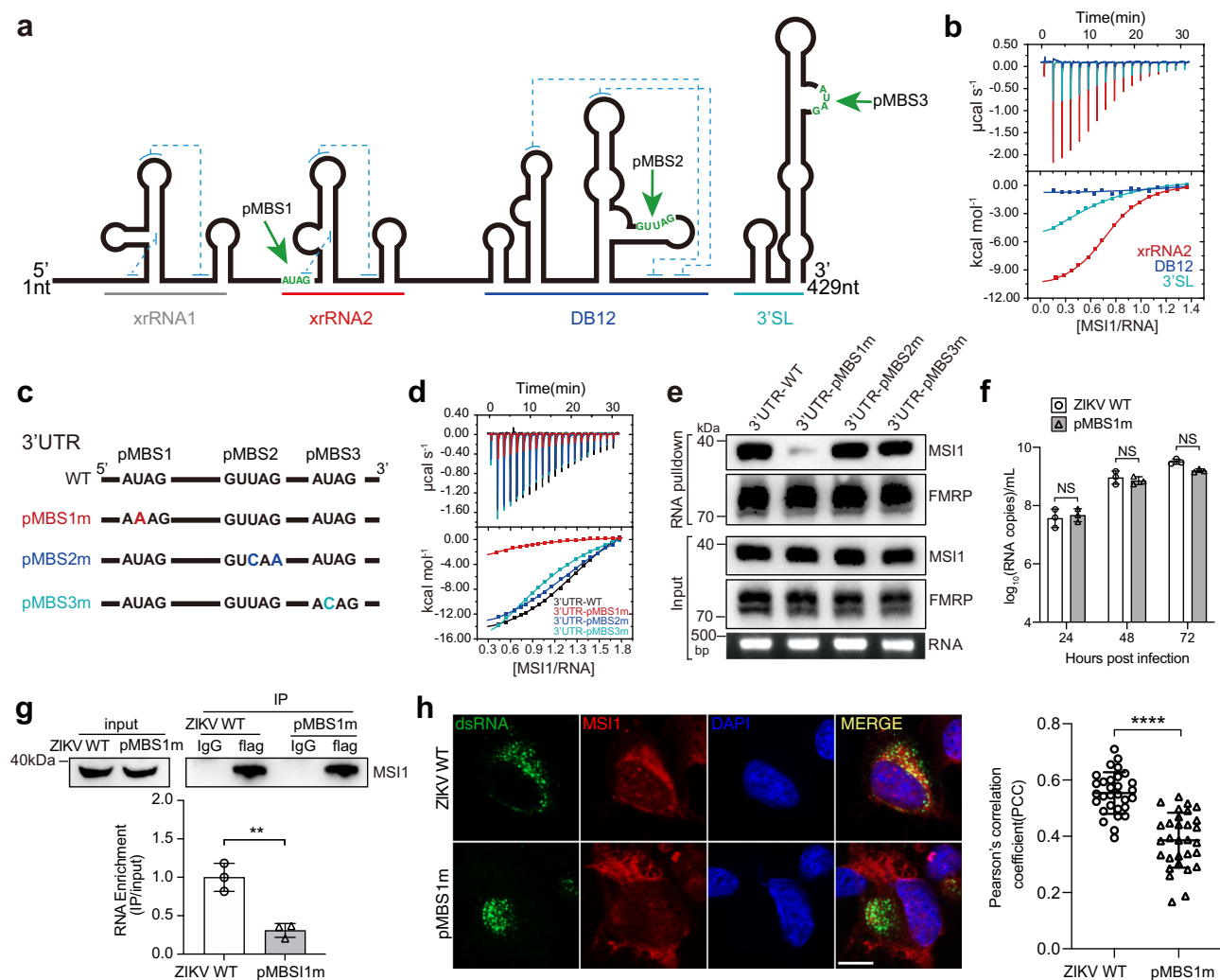
We further sought to determine the function of pMBS1 during ZIKV replication. A panel of human cell lines that endogenously express MSII (Fig. S2e) was subjected to infection with ZIKV and the corresponding pMBS1 mutant virus. Compared with the WT ZIKV, the replication of pMBS1m was significantly decreased over the course of infection in hNPCs (Fig. 2f, g). Similarly, the pMBS1m mutant virus replicated to lower levels compared to the WT virus in U251 and SH-SY5Y cells (Fig. S2c, d). These results showed pMBS1 promoted ZIKV replication in a cell-specific manner.

A BHK-21-MSII cell line with stable MSII expression and the corresponding control cell line (BHK-21-Ctrl) were generated by lentivirus transduction (Fig. S2e) and the replication kinetics of the WT and pMBS1m mutant virus were compared. Similarly to previous results (Fig. 1f), the pMBS1m and WT viruses showed similar replication efficiency in BHK-21-Ctrl cells lacking endogenous MSII. In contrast, the growth of the pMBS1m virus was significantly decreased, compared with the WT virus in BHK-21-MSII cells that stably express MSII (Fig. 2c and Fig. S2h). Furthermore, knockdown of MSII in hNPCs significantly down-regulated ZIKV E protein expression (Fig. 2d) as well as viral RNA amounts (Fig. 2e) after WT ZIKV infection, whereas there was no impact on the growth of pMBS1m.

The pMBS1 mutation was also introduced into a ZIKV replicon (ZIKVrep) containing a *Renilla* luciferase reporter gene (Fig. S2i), resulting in the mutant ZIKV replicon Rep-pMBS1m. Rep-pMBS1m replicated as efficiently as WT in BHK-21-Ctrl cells but with less efficiency in BHK-21-MSII cells (Fig. S2j, k). Furthermore, Rep-pMBS1m displayed sharply decreased RNA replication in hNPCs which normally express MSII (Fig. S2l). Collectively, the results demonstrate that pMBS1 positively regulates ZIKV replication in an MSII-dependent and cell-specific manner.

### The AGAA tetraloop fold in xrRNA2 is critical for MSII binding

The structural basis of MSII binding to the identified MBS of ZIKV was then investigated. The structure of ZIKV xrRNA1 has been determined with a ring-like 3D structure<sup>32</sup>. The homology-derived secondary structure of xrRNA2 was predicated to include a three-way junction formed by P1, P2, and P3 and an additional P4 helix (Fig. 3a, left). As the corresponding P2 and P4 structures in ZIKV xrRNA1 were altered to stabilize the construct for crystallization<sup>32</sup>, we also constructed a similar xrRNA2 mutant (xrRNA2 P2M/P4M), in which the apical loops capping the P2 and P4 stems were simultaneously mutated to produce a GAAA tetraloop (Fig. 3a, right). Unexpectedly, the P2M/P4M mutant completely lost the ability to bind MSII (Fig. 3b). To investigate the lack of MSII binding further, two xrRNA2 mutants were constructed in



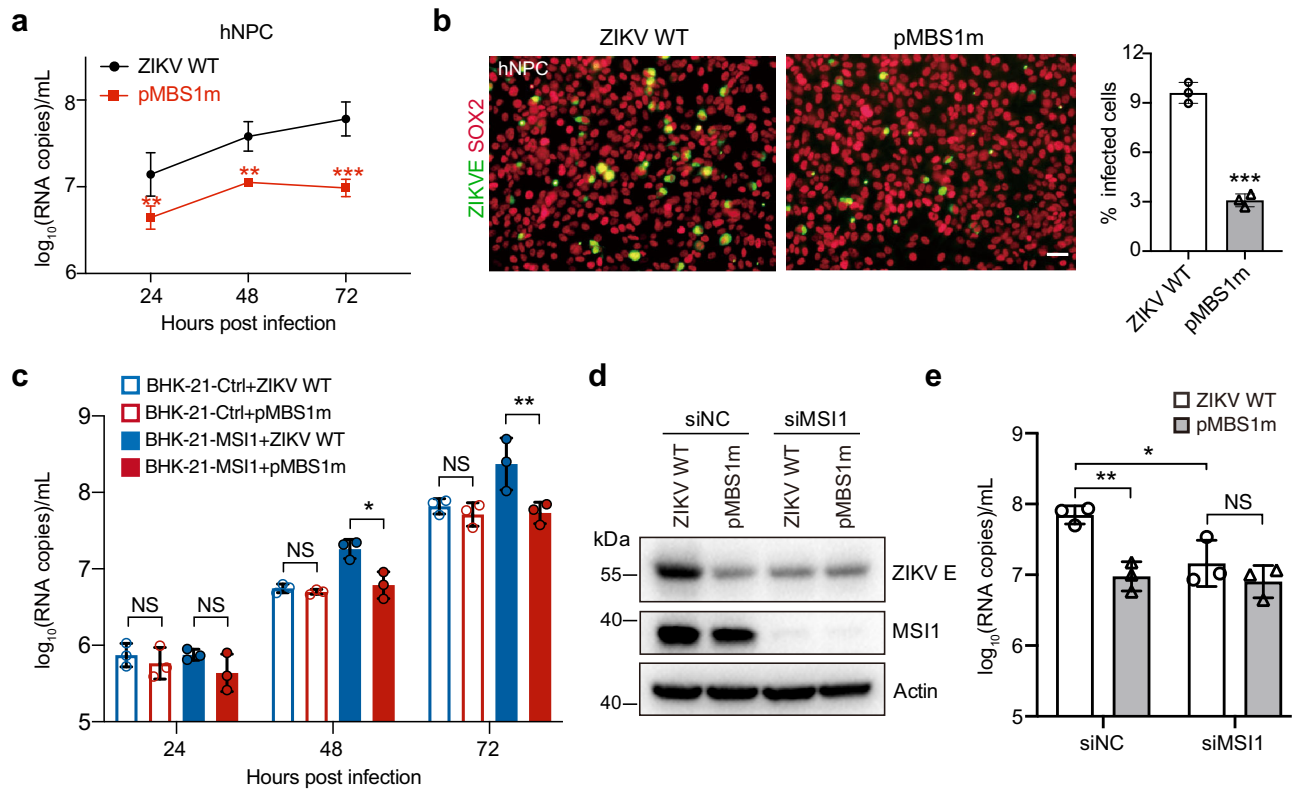
**Fig. 1 | The pMBS1 in ZIKV xrRNA2 is responsible for MSII binding.** **a** Putative MSII binding sites (pMBS) in the 3' UTR of ZIKV. The pseudoknots are indicated with blue dashed lines. **b** The ITC profiles of RRM12 to the individual xrRNA2, DB12, and 3' SL sequences. **c** Schematic showing mutations in the putative MSII binding sites in the 3' UTR of ZIKV. **d** The ITC profiles of RRM12 to the ZIKV 3' UTR WT, pMBS1m, pMBS2m and pMBS3m. **e** RNA pull-down assays performed with the WT or single pMBS mutant (1-3) 3' UTRs of ZIKV. In vitro transcribed biotinylated RNAs were incubated with cell extracts of BHK-21 cells transfected with flag-MSII, and RNA-protein complexes were captured on streptavidin beads. Representative Western blots probed with antibody against MSII are shown together with corresponding protein and RNA inputs. Similar result was repeated independently in 3 times. **f** ZIKV WT and pMBS1m viral RNA copies in culture supernatants of BHK-21 cells (MOI = 0.1). Data are the mean  $\pm$  SD.  $n = 3$  independent experiments. Two-way ANOVA, NS, not significant. **g** RNA-IP analysis from WT or pMBS1m infected BHK-21

cells transfected with flag-MSII. Western blot shows immunoprecipitations (IPs) by immunoglobulin G (IgG) and flag antibodies. Input (5%) represents whole-cell extract. Western blot was probed with antibodies against flag. The bound viral RNA from the IP was analyzed by qRT-PCR. Graph below shows qRT-PCR performed on bound RNA from IP. RNA-IP values are presented as the ratio to the input after subtraction of the IgG background. Data are mean  $\pm$  SD.  $n = 3$  independent experiments. Two-sided Student's  $t$  test,  $**P < 0.01$ . ( $P = 0.0042$ ). **h** Confocal microscopy of BHK-21 cells transduced with MSII-expressing lentiviruses (MOI = 10) and infected with ZIKV WT or pMBS1m (MOI = 1) for 24 h. Fixed cells were immunostained with anti-dsRNA and anti-MSII antibodies. Cell nuclei were counterstained with DAPI. Scale bar, 10  $\mu$ m. Graph on the right displays the Pearson's correlation coefficient (PCC), mean  $\pm$  SD was calculated from 30 cells in each group. Two-sided Student's  $t$  test.  $****P < 0.0001$ . ( $P < 0.0001$ ). Source data are provided as a Source Data file.

which the apical loops capping the P2 or P4 stems were separately replaced with GAAA tetraloops (P2M and P4M) and subjected to MSII binding analysis by ITC (Fig. S3a and Table S2). The P2M mutant lost the capacity to bind MSII which was preserved for the P4M mutant (Fig. 3c), suggesting that the P2 tetraloop, but not P4, is involved in MSII binding. This was further demonstrated by the deletion of the whole P4 ( $\Delta P$ ) structure, which did not disturb MSII-xrRNA2 binding (Fig. S3a, Fig. 3c and Table S2). Additionally, SAXS analysis showed that the overall structure of the P2M mutant is similar to that of the WT xrRNA2 (Fig. S3b-d and Table S1). These results suggested that the apical P2 tetraloop in xrRNA2 functions as a non-canonical MBS.

The AGAA loop ending the P2 stem of xrRNA2 belongs to the AGNN-type tetraloop, in which a non-canonical A-A base pair is formed

closing the helix, and the intermediate guanine and adenine are flipped out<sup>33,34</sup>. The GAAA tetraloop is a typical GNRA-type tetraloop, which also has a non-canonical G-A base pair closing the helix, but the three adenines are stacked with the 3' guanine<sup>34</sup> (Fig. 3d). To clarify whether the tetraloop fold is critical for MSII binding, several xrRNA2 mutants with modified P2 apical loops were constructed and analyzed for their ability to bind MSII (Fig. S3e). The AGNN-type mutants, in which the AGAA loop was replaced with AGCU (P2-AGCU) or AGUU (P2-AGUU), retained strong MSII binding affinities. By contrast, the GNRA-type mutants of xrRNA2, in which the AGAA loop was replaced with GAGA (P2-GAGA) or GCAA (P2-GCAA), showed significantly decreased MSII binding affinities (Fig. 3e and Table S2). These results demonstrated that the P2 tetraloop fold of xrRNA2 is critical for MSII binding.



**Fig. 2 | The pMBS1m mutation attenuates ZIKV replication in cells with MSII expression.** **a** hNPCs were infected with the ZIKV WT or pMBS1m viruses (MOI = 1), and the culture supernatants were harvested at the indicated time points for detection of viral RNA copies by qRT-PCR. Data are mean  $\pm$  SD.  $n = 3$  independent experiments. Two-way ANOVA,  $**P < 0.01$ ,  $***P < 0.001$ . (24 h  $P = 0.0083$ ; 48 h  $P = 0.0054$ ; 72 h  $P = 0.0002$ ). **b** The expression of viral envelope protein at 48 h after infection from (a) was detected by immunostaining. Scale bar, 50  $\mu$ m. Data are mean  $\pm$  SD. Two-sided Student's  $t$  test,  $***P < 0.001$ . ( $P = 0.0001$ ). **c** BHK-21-Ctrl and BHK-21-MSII cells were infected with the ZIKV WT or pMBS1m viruses (MOI = 0.01), and culture supernatant was harvested at the indicated time points for detection of viral RNA copies by qRT-PCR. Data are mean  $\pm$  SD.  $n = 3$  independent experiments.

Two-way ANOVA,  $*P < 0.05$ ,  $**P < 0.01$ , NS not significant. (48 h: BHK-21-MSII-WT vs BHK-21-MSII pMBS1m,  $P = 0.0438$ ; 72 h: BHK-21-MSII-WT vs BHK-21-MSII pMBS1m,  $P = 0.0021$ ). **d, e** hNPCs were treated with negative control siRNA (siNC) and MSII siRNA (siMSII) for 36 h, and then infected with the ZIKV WT or pMBS1m viruses (MOI = 3), and at 48 h after infection, viral envelope and MSII in cell lysates was detected by Western blot (d), culture supernatant were harvested for detection of viral RNA copies by qRT-PCR (e). Data are mean  $\pm$  SD.  $n = 3$  independent experiments. Two-sided Student's  $t$  test,  $*P < 0.05$ ,  $**P < 0.01$ , NS not significant. (siNC WT vs siNC pMBS1m,  $P = 0.0035$ ; siNC WT vs siMSII WT,  $P = 0.0275$ ). Source data are provided as a Source Data file.

The effects of structure-based mutation of P2M on MSII binding in vitro led us to predict that the same mutation would alter MSII binding ability during infection. Therefore, we produced a P2M mutant ZIKV, in which the xrRNA2 P2 was replaced with the sequence 5' CC-GAAA-GG 3' and investigated the effects of the mutation on virus replication. The plaque and sRNA production phenotypes of the WT and P2M viruses were similar (Fig. S3f, g), and both viruses replicated equally well in BHK-21 cells that lacked endogenous MSII expression (Fig. S2e). RNA-IP was performed using lysates prepared from ZIKV WT or P2M infected BHK-21 cells transfected with MSII, which revealed a significantly reduced interaction between MSII and the P2M RNA compared to the WT virus RNA (Fig. 4a); While the interaction of the control protein FMRP with P2M RNA is similar to that of WT virus RNA (Fig. S3i). Consistently, there was less colocalization observed between MSII and double-stranded RNA in ZIKV P2M-infected cells compared to WT virus-infected cells (Fig. 4b). These data confirmed the important role of the xrRNA2 P2 tetraloop fold in mediating in vitro and in vivo MSII binding.

More importantly, the P2M mutation significantly reduced virus propagation and replicon RNA synthesis in BHK-21-MSII cells compared to the WT equivalents, whereas no differences were observed in BHK-21-Ctrl cells (Fig. S3g–j). Furthermore, in hNPCs, ZIKV P2M showed markedly lower levels of viral RNA and E protein expression compared with WT virus (Fig. 4c, d), and the Rep-P2M replicon

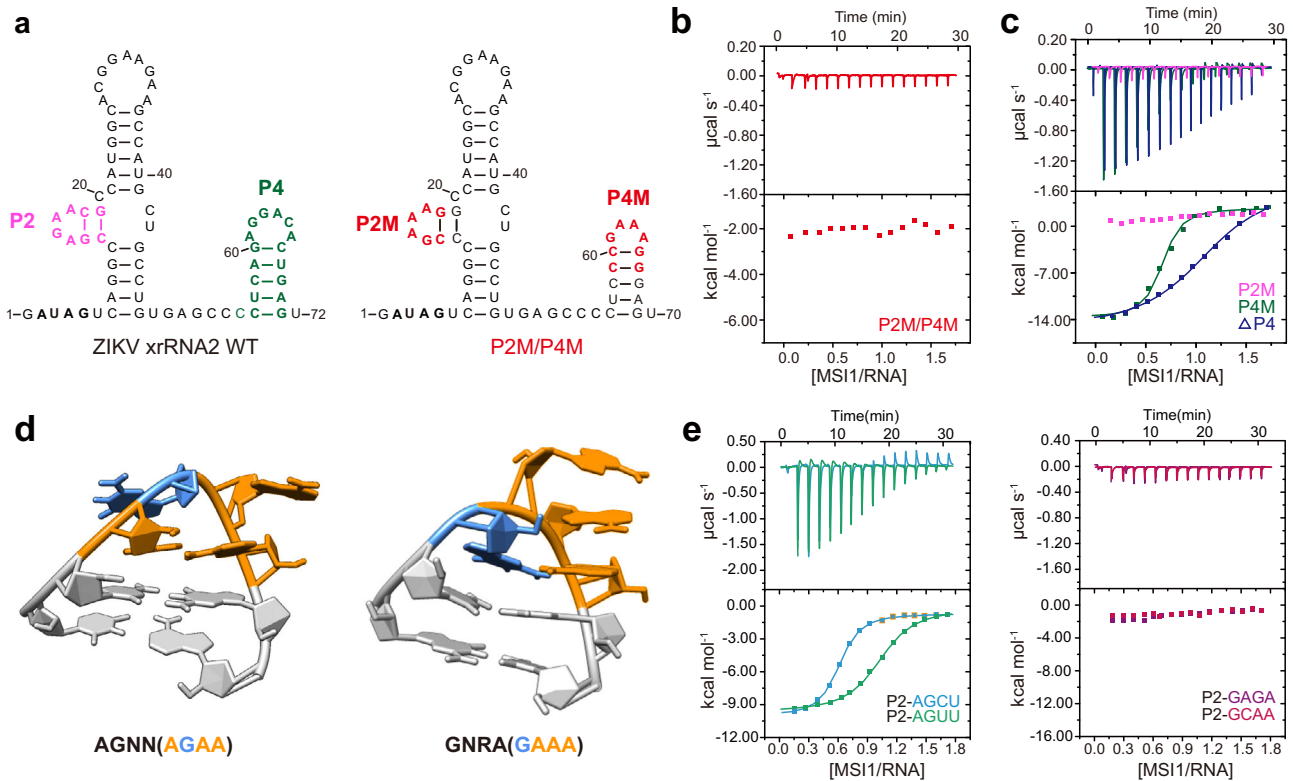
produced much less luciferase signal than the WT replicon at 48 h post-transfection (Fig. S3k). Collectively, these data demonstrated that the AGAA tetraloop of xrRNA2, like pMBS1, regulates ZIKV replication in an MSII-dependent manner, representing a non-canonical MBS.

To rule out the possibility that attenuation of P2M ZIKV was owing to a non-specific effect of P2 sequence change, we produced an additional mutant ZIKV, P2-AGCU (Fig. S4a), in which the xrRNA2 P2 was replaced with the sequence 5' CG-AGCU-CG 3', keeping the AGNN-type of P2 unchanged. P2-AGCU ZIKV and WT ZIKV replicated equally well in BHK-21 cells and hNPCs (Fig. S4b–d), which was consistent with that the P2-AGCU mutation did not affect MSII binding affinity (Fig. 3e). These data reinforced the conclusion that the contribution of xrRNA2 P2 to ZIKV replication depends on its AGNN-type structure.

#### The AGNN-type tetraloop of xrRNA2 is only present in ZIKV

Considering the integral role of the identified MBSs in the ZIKV genome, we determined whether similar MBSs were also present in the 3' UTRs of other viruses by sequence alignments and RNA structure predictions. Intriguingly, a large number of viruses, including coxsackievirus B3 (CVB3), hepatitis C virus (HCV), human rhinovirus 14 (HRV14), Norwalk virus (NV), western equine encephalitis virus (WEEV), Sindbis virus (SINV), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and murine hepatitis virus (MHV), contain pMBS sequences, with SARS-CoV-2 and MHV containing





**Fig. 3 | xrRNA2 P2 tetraloop fold is critical for MSII binding.** **a** Secondary structure of ZIKV xrRNA2 and the xrRNA2 mutant P2M/P4M. **b** ITC profiles of RRM12 to xrRNA2 P2/P4M. **c** The ITC profiles of RRM12 to xrRNA2 mutants P2M, P4M and  $\Delta$ P4. **d** NMR structures of tetraloops of AGNN and GNRA. **e** ITC profiles of

RRM12 to xrRNA2 P2 loop mutants of ZIKV xrRNA2. Mutants with AGNN-type tetraloop include P2-AGCU and P2-AGUU. Mutants with a GNRA-type tetraloop include P2-GAGA and P2-GCAA.

seven and five pMBSs in their 3' UTRs respectively (Fig. S5a). As for flaviviruses, the vast majority of them contain the pMBS sequence (Fig. S5b), while none of these viruses harbor the A/GUAG and AGNN-type tetraloop simultaneously, except for ZIKV (Fig. 5b). Some flaviviruses contain the same AUAG motif in their xrRNAs, but their P2 loops don't belong to the AGNN-type. Further analysis of MSII binding showed that the 3' UTRs from non-ZIKV flaviviruses including MVEV, USUV, JEV, YFV, and DENV exhibited no binding affinity to MSII (Fig. 5c). We then tested the impact of MSII on the replication of these (mosquito-borne flaviviruses) MBFVs. As expected, MSII knockdown by RNAi in hNPCs reduced ZIKV replication, while the replication of YFV, JEV, WNV, DENV2, and DENV4, were not affected by MSII knockdown (Fig. 5d and e). More importantly, phylogenetic analysis revealed that ZIKV 3' UTRs clustered in distinct clades: American, Asian, and African, but the AUAG motif, as well as the AGAA tetraloop, are both evolutionally conserved in all ZIKV isolates (Fig. S6). Collectively, these results indicated that the unique AGAA tetraloop of xrRNA2 distinguishes ZIKV from other flaviviruses.

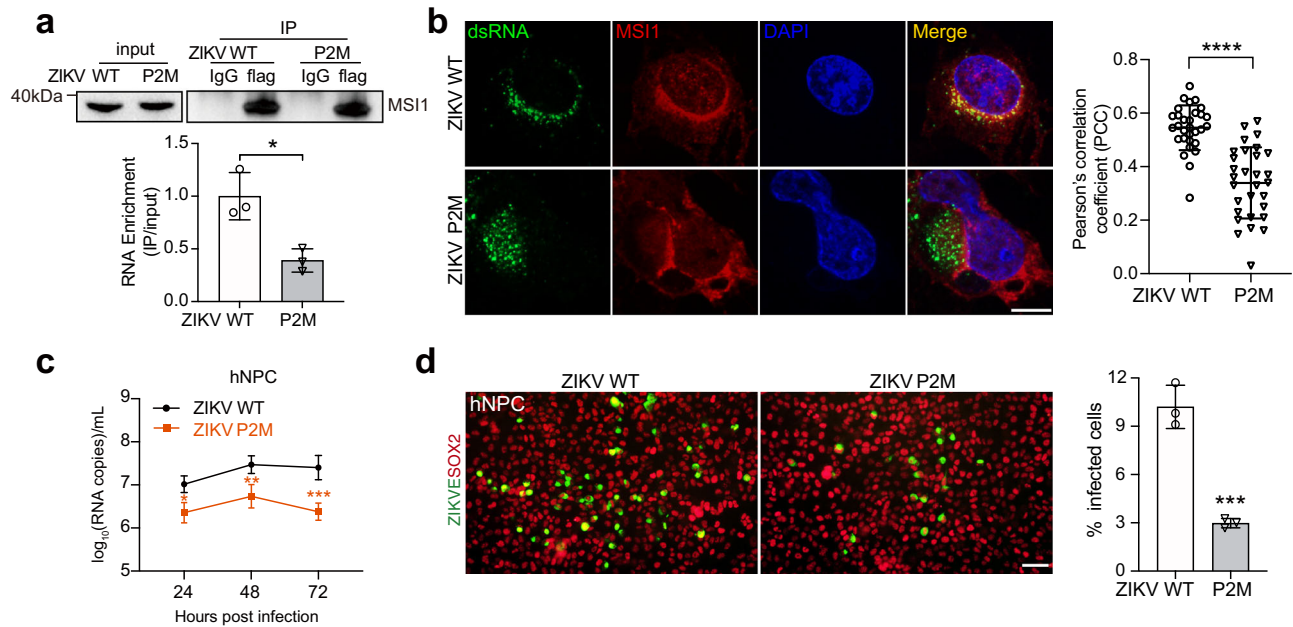
### The MSII binding element of ZIKV xrRNA2 confers MSII binding ability to other flaviviruses

We next investigated whether the MSII binding element alone is sufficient to endow other flaviviruses neurotropism like ZIKV. As shown above, DENV4 xrRNA does not contain A/GUAG sequence nor AGNN-type tetraloop (Fig. 5b), and has no MSII binding ability (Fig. 5c). We then transplanted the pMBS sequence and AGAA tetraloop of ZIKV xrRNA2 to the corresponding positions of DENV4 xrRNA. As shown in Fig. 6A, the A<sub>3</sub> and A<sub>13</sub>C<sub>14</sub>U<sub>15</sub>U<sub>16</sub>G<sub>17</sub>U<sub>18</sub> of DENV4 xrRNA were mutated into U<sub>3</sub> and G<sub>13</sub>A<sub>14</sub>G<sub>15</sub>A<sub>16</sub>A<sub>17</sub>C<sub>18</sub>, respectively, generating a DENV4 xrRNA mutant (DxM), to endow it with similar configuration to ZIKV xrRNA2. Then we detected the binding between DxM RNA and MSII by

ITC assay. MSII shows high affinity to DxM xrRNA with K<sub>d</sub> at 2.5  $\mu$ M, which is similar to that of ZIKV xrRNA2 (Fig. 6b, left). The binding between DENV4 3'UTR mutant and MSII is also enhanced (Fig. 6b, right). We subsequently introduced this mutant to a DENV4 814669 strain infectious clone<sup>24</sup>, and constructed an xrRNA mutated DENV4 virus (DENV4 DxM). BHK-21 cells or hNPCs were infected with WT DENV4 or DxM and analyzed viral replication. DxM infection measured by viral RNA replication was similar to WT in BHK-21 cells (Fig. 6c). More importantly, DxM exhibited enhanced replication ability in hNPCs (Fig. 6d, e). Thus, transplantation of MSII binding element of ZIKV xrRNA2 could confer other flaviviruses enhanced replication in MSII expressing cells.

### MSII RRM12 utilize two different RNA binding interfaces to bind to xrRNA2

The two RRMs of MSII adopts a canonical  $\beta$ 1- $\alpha$ 1- $\beta$ 2- $\beta$ 3- $\alpha$ 2- $\beta$ 4 topology, and the conserved aromatic side chains (termed the RNP consensus sequence) that are essential for RNA binding are located on the  $\beta$ 1 (RNP2) and  $\beta$ 3 (RNP1) strands. However, the AGAA tetraloop doesn't belong to the conserved MSII binding sequence, and how the two homogeneous RRMs bind to two different RNA sequences is still unknown. To explore whether MSII utilizes the AGAA tetraloop with the canonical RNPs, we determined the interaction interfaces within RRM12 using Hydrogen-Deuterium Exchange mass spectrometry (HDX-MS). Experiments were performed for RRM12 alone (*apo*-form) or in the presence of a 2-fold molar excess of xrRNA2 (*holo*-form) (Table S5). A total of 98 and 102 peptides from the *apo*-form and *holo*-form RRM12, respectively, were identified and characterized across all HDX time points, representing a high sequence coverage (95%) (Fig. S7a, Supplementary Data 1). In the *apo* state, most peptides exhibited high solvent accessibility with relative deuterium uptake



**Fig. 4 | The xrRNA2 AGAA tetraloop participates in MSII binding in vivo and regulates ZIKV replication.** **a** RNA-IP analysis from WT or P2M infected BHK-21 cells transfected with flag-MSII. Western blot shows immunoprecipitations (IPs) by immunoglobulin G (IgG) and flag antibodies. Input (5%) represents whole-cell extract. Western blot was probed with antibodies against flag. The bound viral RNA from the IP was analyzed by qRT-PCR. Graph below shows qRT-PCR performed on bound RNA from IP. RNA-IP values are presented as the ratio to the input after subtraction of the IgG background. Data are mean  $\pm$  SD.  $n = 3$  independent experiments. Two-sided Student's  $t$  test,  $*P < 0.05$ . ( $P = 0.0237$ ). **b** Confocal microscopy of BHK-21 cells transfected with MSII-expressing lentiviruses (MOI = 10) and infected with the ZIKV WT or P2M viruses (MOI = 1) for 24 h. Fixed cells were

immunostained with anti-dsRNA and anti-MSII antibodies. Cell nuclei were counterstained with DAPI. Scale bar, 10  $\mu$ m. The graph on the right shows Pearson's correlation coefficient (PCC), and mean  $\pm$  SD was calculated from 30 cells in each group. Two-sided Student's  $t$  test.  $****P < 0.0001$ . ( $P < 0.0001$ ). **c** hNPCs were infected with the ZIKV WT or P2M viruses (MOI = 1), and the culture supernatants were harvested at the indicated time points for detection of viral RNA copies by qRT-PCR. Data are mean  $\pm$  SD.  $n = 3$  independent experiments. Two-way ANOVA,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ . (24 h  $P = 0.0139$ ; 48 h  $P = 0.0069$ ; 72 h  $P = 0.0005$ ). **d** The expression of viral envelope protein at 48 h after infection from (c) was detected by immunostaining. Scale bar, 50  $\mu$ m. mean  $\pm$  SD. Two-sided Student's  $t$  test,  $***P < 0.001$ . ( $P = 0.0008$ ). Source data are provided as a Source Data file.

fractions equal to or above 60% after 5 minutes incubation, suggesting RRM12 has a highly dynamic conformation in solution (Fig. S7b). Both RRM1 and RRM2 adopt a  $\beta\alpha\beta\beta\alpha\beta$  fold topology<sup>16,19,35,36</sup>. Four peptides corresponding to residues 35–42 ( $\alpha_1$ ) and 76–85 ( $\alpha_2$ ) of RRM1 and residues 107–120 ( $\beta_1$ ) and 142–157 ( $\beta_3$ ) of RRM2 showed substantial decrease in HDX kinetics in the *holo* state (Fig. S7b, c), implying that they may be at or close to the binding interfaces (Fig. 7b). Canonical RRM-RNA binding interactions indicate that conserved residues in the  $\beta_1$  and  $\beta_3$  strands of RRM2 are involved in single-stranded RNA recognition, while residues at helices  $\alpha_1$  and  $\alpha_2$  are not<sup>16,19,35,36</sup>. The HDX-MS data indicate that RRM2 uses the canonical RNA recognition interfaces whilst RRM1 may utilize a non-canonical interface.

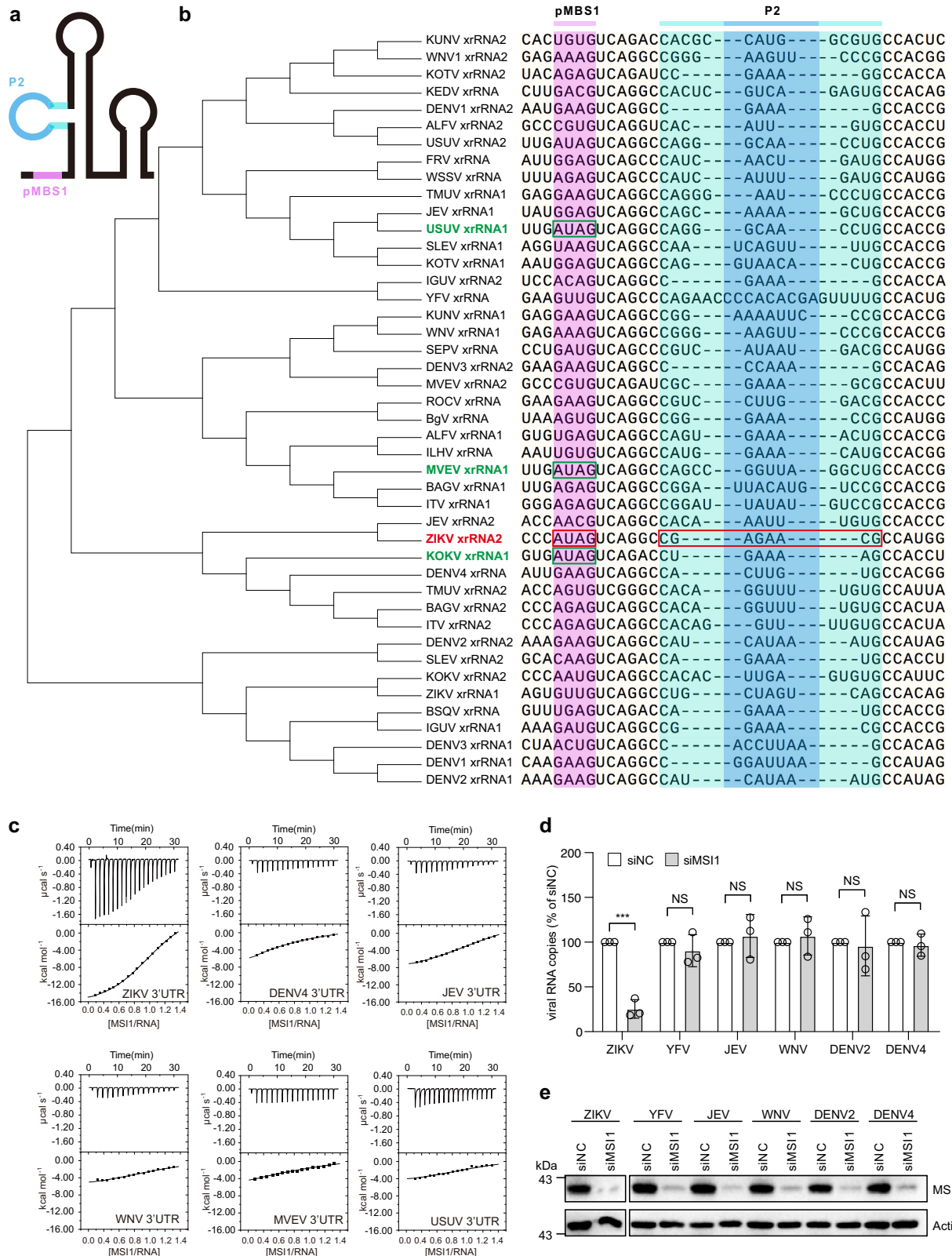
To validate the predicted RRM1-RNA interface, we monitored the effect of alanine substitutions of residues residing in the non-canonical interface identified by HDX-MS ( $\alpha_1$ : R37A/F40A/F43A,  $\alpha_2$ : K76A/R82A) or the canonical RNA recognition interface ( $\beta_1$ : K21A/F23A,  $\beta_3$ : F63A/F65A) in RRM1 by ITC experiments. SAXS data demonstrated that the described mutations didn't cause significant structural changes to RRM12 (Fig. S7e–g and Table S1). While substitutions of residues in the canonical interface ( $\beta_1$  and  $\beta_3$ ) of RRM1 did not impair xrRNA2 binding, substitutions of residues in  $\alpha_1$  and  $\alpha_2$  attenuated xrRNA2 binding (Fig. S7h, Table S2). Thus, RRM1 utilizes a non-canonical interface at or close to helices  $\alpha_1$  and  $\alpha_2$  to bind xrRNA2.

### Structural investigation of the MSII-xrRNA2 complex by integrative methods

To investigate the potential binding model of MSII in complex with xrRNA2, an integrative structural approach that combines SAXS, interface information from mutagenesis and HDX-MS experiments with computational modeling was utilized<sup>37</sup>. We first characterized the

solution structure of the MSII RRM12 in complex with xrRNA2 at low resolution by SAXS (Fig. 7). The overall 3D shape envelopes for RRM12, xrRNA2 and the RRM12-xrRNA2 complex were *ab initio* reconstructed from the SAXS data<sup>38</sup> (Fig. 7a–e). Both the atomic models of MSII and xrRNA2 by SAXS-driven homology modeling can be nicely fitted into the corresponding 3D shape envelopes by DAMMIN (Fig. 7b, c). The two-phase modeling of the RRM12-xrRNA2 complex by MONSA<sup>38</sup> suggests a side-by-side interaction between the protein and RNA (Fig. 7d).

We initially performed *de novo* docking of RRM12 and xrRNA2 complex without experimental restraints using HADDOCK2.4 (high ambiguity driven docking) platform<sup>39,40</sup>. The generated models fit SAXS data badly ( $\chi^2$  greater than 10) and the binding interfaces are conflicting with experimental data, such as RRM12 binding to P4 or PK2 of xrRNA2. We next performed information-driven flexible docking on the RRM12-xrRNA2 complex. The binding interface restraints from mutagenesis and HDX-MS data as well as the radius of gyration ( $R_g$ ) restraint from SAXS data were supplemented during integrative modeling by HADDOCK 2.4. HADDOCK clustered 230 structures in 10 clusters, with Z-scores ranging from  $-1.4$  to  $1.3$  (the smaller the better). The experimental SAXS scattering profile and the 3D shape envelope by DAMMIN were used to validate the models. As shown in Fig. 7e–g, the representative models from the top 2 best clusters (Z-scores of  $-1.4$  and  $-1.2$ , respectively) fit well with the experimental SAXS data (fitting  $\chi^2$  of 2.5 and 3.2, respectively) and the shape envelope by DAMMIN. The binding interfaces between xrRNA2 and RRM12 in these models are also compatible with the mutagenesis and HDX-MS data. The models from clusters 3–10 either fit the SAXS data poorly (fitting  $\chi^2$  larger than 10), or are not compatible with the binding interface data (Fig. S8), are therefore not further discussed. The representative



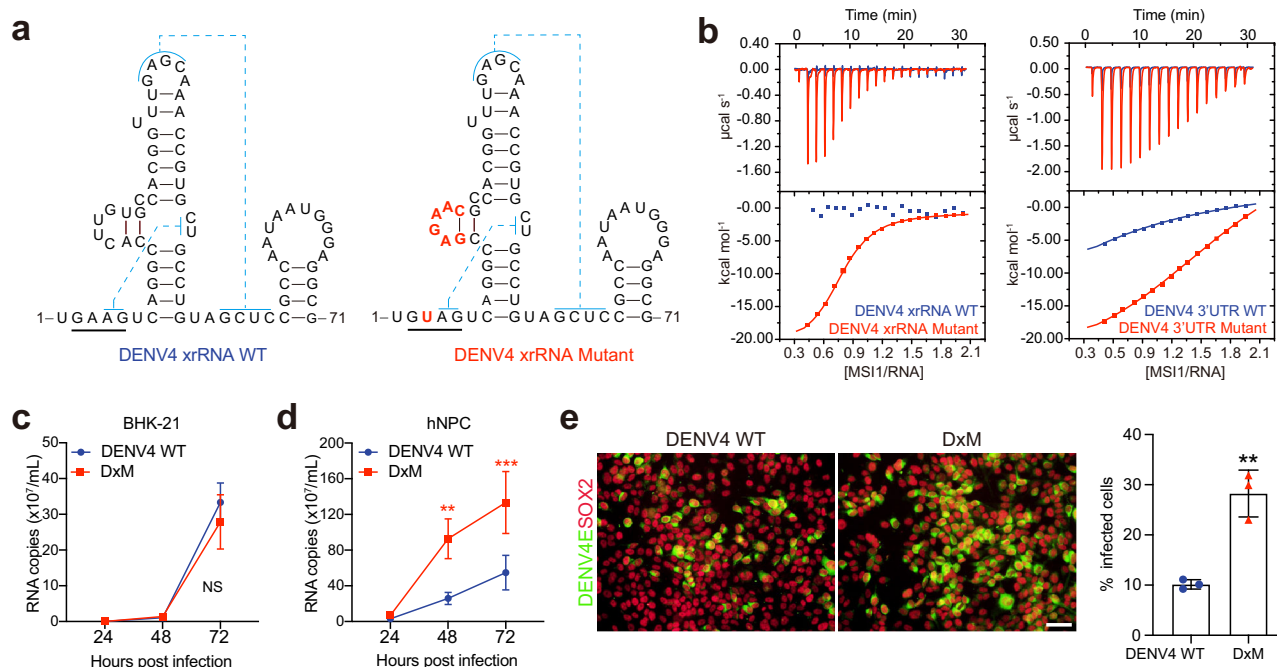
model from cluster 1 (Z-score of  $-1.64$ ), as shown in Fig. 7f, reveals a bipartite binding mode, which the pMBS1 and P2 AGNN tetraloop of xrRNA2 act like a pincer to specifically target the noncanonical RNA binding interface in RRM1 ( $\alpha 1$  and  $\alpha 2$ ) and loop3 connecting the  $\beta 2$  and  $\beta 3$  strands in RRM2 of MS1, respectively. These binding interfaces between xrRNA2 and RRM1/2 are also compatible with the HDX-MS and mutagenesis data. In the representative model (Z-score of

$1.2$ ), RRM2 is assumed to utilize the canonical RNA binding interfaces of the  $\beta 1$  (RNP2) and  $\beta 3$  (RNP1) strands as well as loop3 (connecting  $\beta 2$  and  $\beta 3$ ) to recognize the pMBS1 (U3A4G5), and RRM1 recognizes the AGNN-type P2 of xrRNA2 through its  $\alpha$ -helices (Fig. 7g). The interaction of RRM1 with the AGAA tetraloop in P2 is reminiscent of the recognition of the AGNN tetraloop by the double-stranded RNA binding domain (dsRBD) of Rnt1p RNase III33, where the dsRBD



**Fig. 5 | The two MBSs of xrRNA2 are unique among MBFV.** **a** Secondary structure diagram of flaviviral xrRNA with the pMBS1 colored in pink, the P2 loop colored in blue and the P2 stem colored in cyan. **b** A maximum likelihood tree constructed with the xrRNA sequences of flaviviruses using MEGA. The pMBS1 and P2 loop sequences (colored as in **a**) of xrRNAs are shown after the virus names. The pMBS1 and P2 loop sequences of ZIKV xrRNA2 are framed with red rectangles. The pMBS1 sequences of the xrRNAs of KOKV, USUV, and MVEV, which match AUAG are framed with green rectangles. The GenBank ID information of sequences used for this analysis were listed in Table S4. **c** ITC profiles of RRM12 to JEV/YFV/DENV2/

DENV4/WNV 3' UTRs. **d** Relative viral RNA copies in the supernatants of control siRNA (siNC) and MSII siRNA (siMSII) treated hNPCs after infection with different flaviviruses (ZIKV: MOI = 1; YFV: MOI = 1; JEV: MOI = 1; WNV: MOI = 0.1; DENV2: MOI = 1; DENV4: MOI = 1). **e** Representative Western blots of hNPCs treated with control and MSII siRNAs. Blots were probed with antibodies against MSII or actin as a loading control. Data were expressed as the mean  $\pm$  SD.  $n = 3$  independent experiments. Two-sided Student's  $t$  test, \*\*\* $P < 0.0001$ , NS not significant. (ZIKV  $P = 0.0002$ ). Source data are provided as a Source Data file.



virus (MOI = 1), and the culture supernatants were harvested at the indicated time points for detection of viral RNA copies by qRT-PCR. Data are mean  $\pm$  SD.  $n = 3$  independent experiments. Two-way ANOVA, \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . (48 h  $P = 0.0028$ ; 72 h  $P = 0.0008$ ). **e** The expression of viral envelope protein at 48 h after infection from (d) was detected by immunostaining. Scale bar, 50  $\mu$ m. mean  $\pm$  SD.  $n = 3$  independent experiments. Two-sided Student's  $t$  test, \*\* $P < 0.01$ . ( $P = 0.0027$ ). Source data are provided as a Source Data file.

contacts the RNA at successive minor, major, and AGNN tetraloop minor grooves on one face of the helix. Taken together, the models in cluster 2 also indicated a bipartite binding mode between MSII and xrRNA2.

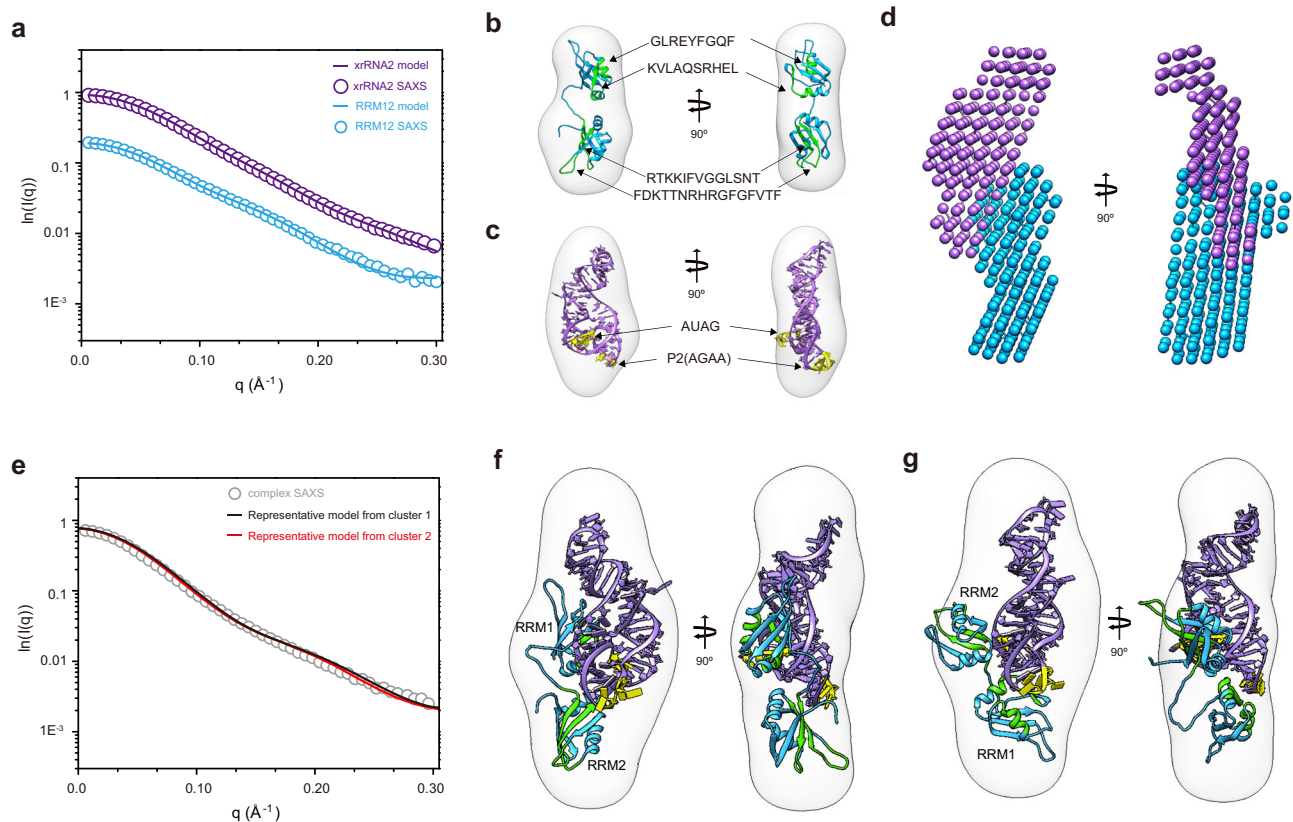
## Discussion

Here, we demonstrated ZIKV utilizes its canonical pMBS1 (5'-AUAG-3') and a novel non-canonical MBS, the AGNN-type tetraloop in xrRNA2 to specifically target the RRM1 and RRM2 subdomains of MSII in a bipartite mode. Of the three pMBSs in the ZIKV 3'-UTR, only pMBS1 in xrRNA2 (5'-AUAG-3') is critical for MSII binding and viral replication in hNPCs. Thus, the AUAG motif represents the authentic canonical MBS, which is well documented in host RNA transcripts. Surprisingly, we also revealed that the AGAA tetraloop of xrRNA2 is also critical for MSII binding and the unique cell tropism of ZIKV, thus representing a second but non-canonical MBS. More importantly, the MSII binding affinity is completely dependent on the cooperation of the two MBSs.

Our findings have significant implications. First, the identification of the AGNN-type tetraloop as an MBS changed the canonical rule of MSII binding targets. All previous RNA targets of MSII are identified as

an A/GU<sub>(1-3)</sub>AG motif<sup>11,13,18,21,22</sup>, this long-standing information may be incomplete. Our findings contribute to a better understanding of the complexity of MSII binding. Whether similar structures are present in host RNA transcripts remains to be determined. This unexpected binding model of ZIKV xrRNA2 with MSII suggests the known MSII targets of host RNA transcripts, such as numb<sup>11</sup> and p21<sup>20</sup>, may be more complex and include unidentified RNA structures. Second, our findings provide a long-sought answer to the question of why is only ZIKV linked to neurodevelopmental diseases. A wide variety of viruses possess the canonical pMBS sequence in their 3' UTRs<sup>26</sup> (Fig. S5, Fig. 5a), but none of them has been proven to have the ability to bind MSII. The unique MBS of ZIKV thus provides new evidence supporting the causal link between ZIKV infection and the distinctive congenital malformation in infants. Finally, our results highlight a new role of xrRNA2. The only known function of the highly structured xrRNA in flaviviruses was to block the degradative activity of host exonuclease Xrn-1 resulting in the generation of sRNAs<sup>32,41</sup>, which facilitate viral replication, by multiple actions including blocking the host IFN response<sup>42,43</sup>, suppressing RNAi<sup>44,45</sup> and apoptosis<sup>46</sup>. Our findings suggest that, in addition to producing sRNAs, ZIKV xrRNA2 can also use





**Fig. 7 | Structural investigation of the MSII RRM12-xrRNA2 complex by integrative methods.** **a** Structural characterization of free RRM12 and free xrRNA2 by SAXS. The back-calculated scattering profiles (solid line) of the best models were overlaid with the respective experimental profiles (open cycle). **b** The homology model of RRM12 (blue) was further subjected to rigid body refinement against SAXS data by Xplor-NIH and fitted into the envelope ab initio. The residues presumably involved in xrRNA2 binding identified by HDX-MS are highlighted in green. **c** The best homology model of xrRNA2 (purple) was fitted into the envelope ab initio. The pMBS1 and AGAA tetraloop of xrRNA2 are highlighted in yellow. **d** The 3D shape

envelope of RRM12-xrRNA2 complex reconstructed ab initio with the multiphase modeling program MONSA, and the bead models for RRM12 and xrRNA2 are colored according to (b) and (c). **e** Structural characterization of RRM12-xrRNA2 complex by SAXS. The back-calculated scattering profiles (solid line) of the representative models from cluster 1 (black) and cluster 2 (red) are overlaid with the respective experimental profile (open cycle). **f, g** The representative model from cluster 1 (f) and cluster 2 (g) are fitted into the 3D shape envelope ab initio reconstructed by DAMMIN. RRM12 and xrRNA2 are colored as in (b) and (c).

its unique RNA motif and tetraloop fold to target MSII which determines viral replication in specific cells, like NPCs.

The structural study of RNA-protein complexes using conventional high-resolution techniques such as X-ray crystallography (XRC), cryo-EM and NMR remains challenging, and currently only about 2.2% of the structures in the Protein Data Bank represents protein-RNA complexes (PDB)<sup>47</sup>. Using an integrative approach combining SAXS, HDX-MS, mutagenesis, and computational modeling, we presented two potential bipartite binding models for hMSII RRM12 and ZIKV xrRNA2, which precisely explain the results obtained from biochemical and virological assays. Given that MSII utilizes its two RRMs to interact with pMBS1 and the P2 stem-loop of xrRNA2, two possible binding modes of RRM12 with xrRNA2 can be speculated: 1) The pMBS1 of xrRNA2 binds to RRM2 and the AGNN-type P2 stem of xrRNA2 binds to RRM1 (mode 1); 2) The pMBS1 of xrRNA2 binds to RRM1 and the AGNN-type P2 stem binds to RRM2 (mode 2). Our HDX-MS data indicated that RRM1 interacts with xrRNA2 with a noncanonical interface ( $\alpha 1$  and  $\alpha 2$ ), while RRM2 interacts with xrRNA2 with the conserved canonical interface ( $\beta 1$  and  $\beta 2$ , as well as loop3 that connects  $\beta 1$  and  $\beta 2$ ). The NMR structures of mouse MSII RRM1 and RRM2 in complex with single-stranded RNA (AUAG, pMBS1) have shown that both RRMs utilize the canonical RNA binding motif to interact with single-stranded RNA<sup>16,19</sup>. The interaction of AGNN tetraloop with  $\alpha$ -helix in proteins has also been reported<sup>33,48</sup>. Thus the binding mode 1 in models from cluster 2 contains more RNA-protein interaction features similar to that in the

published structural data. However, it's currently difficult to ascertain which binding mode is accurate due to the low resolution of SAXS data. A high-resolution structure of the MSII-xrRNA2 complex should be pursued in the future.

Overall, our study not only identifies an unusual viral RNA element that contributes to the unique tropism of ZIKV, but also illustrates the precise mechanism and structural basis by which a virus utilizes a primary RNA sequence motif and a tertiary RNA structure to target a specific host protein (Fig. S9).

## Methods

### Ethics statement

All virus studies were performed in strict accordance with the guidelines set by the Academy of Military Medical Sciences (AMMS). Approval to use JEV vaccine strain SA14-14-2 was obtained from the Chengdu Institute of Biological Products. Approval to use YFV vaccine strain 17D was obtained from the Beijing Institute of Biological Products. Informed consent was obtained during the isolation of ZIKV and DENV strains.

### Cells and viruses

The baby hamster kidney fibroblast (BHK-21) cells (ATCC CCL10), African green monkey kidney epithelial (Vero) cells (ATCC CCL81) were cultured in DMEM (Thermo Fisher Scientific) containing 7% fetal bovine serum (FBS, Biowest). C6/36 cells (ATCC CRL-1660) were

cultured in PRMI 1640 (Thermo Fisher Scientific) containing 10% FBS (Biowest). Human Neuroblastoma cell line SH-SY5Y (ATCC CRL-2266) was cultured using DMEM/F12 (Thermo Fisher Scientific) containing 10% FBS. Human neural progenitor cell (hNPC) line 15167 derived from fetal brains (Lonza) was kindly provided by S. Bao (Cleveland Clinic). hNPC and U251 were cultured as neurospheres in neurocult-XF basal medium (STEMCELL technologies) supplemented with neurocult-XF proliferation supplement (STEMCELL technologies), basic fibroblast growth factor (bFGF, 10 ng/mL, STEMCELL technologies), and epidermal growth factor (EGF, 20 ng/mL, STEMCELL technologies), Heparin solution (2 µg/mL, STEMCELL technologies). For the BHK-21-Ms11 cells stably expressing MS11, the MS11 encoding sequence (NM\_002442.4) was constructed into lentiviral vector LV-6, containing a puromycin resistance gene by Genepharma (Shanghai, China). BHK-21 cells were infected with lentivirus with an MOI of 100. Forty-eight hours after infection, cells were screened with 1 µg/mL puromycin for a week. MS11 expression was detected by IFA and Western Blot. BHK-21-Ctrl cells were constructed by infection with LV-6 lentiviral empty vector.

ZIKV strain FSS13025 (GenBank accession number KU955593) was originally isolated from a patient in Cambodia in 2010<sup>31</sup>. The JEV vaccine strain SA14-14-2 and YFV vaccine strain 17D were from the Chengdu Institute of Biological Products and Beijing Institute of Biological Products, respectively. WNV strain NY99 was rescued from the two-plasmid infectious clone<sup>49</sup>. DENV2 strain 43 (GenBank accession number AF204178.1) was isolated from dengue fever (DF) patient in China in 1987. DENV4 strain B5 (GenBank accession number AF289029.1) was isolated from a DF patient in China<sup>50</sup>.

### Antibodies

MS1-1 (rabbit, ab52865, Abcam), SOX2 (rabbit, ab97959, Abcam), Flag (mouse-F1804, sigma), dsRNA-J2 (mouse, 1010200, Scicons), ZIKV envelope protein (mouse, BF-1176-46, Biofront technologies), FMRP (rabbit, A4539, abclonal), DENV envelope protein (2A10, mouse, produced in our lab), Actin (rabbit, AC026, Abclonal), mouse IgG isotype control (ab37355, Abcam), goat anti-rabbit IgG (alexa fluor 594) (ab150080, Abcam), goat anti-mouse IgG (alexa fluor 488) (ab150113, Abcam), horseradish peroxidase-conjugated goat anti-rabbit IgG (ZB-2301, ZSGB-BIO), horseradish peroxidase-conjugated goat anti-rabbit IgG (ZB-2305, ZSGB-BIO).

### Protein expression and purification

DNA encoding human MS11 RRM12 (20-191), RRM1 (20-103), and RRM2 (109-191) were amplified from the human Musashi-1 gene by polymerase chain reaction (PCR) and subcloned into pGEX-6p-1. MS11 RRM12 cysteine mutants (C49/C167/C181S (for double labeling), C49/C167S/C181S (for single labeling) and C49S/C167/C181S (for single labeling)) for EPR experiment were generated by the Quick-change site-directed mutagenesis strategy and verified by DNA sequencing. The information on primers used was provided in Table S3. The proteins were expressed with an N-terminal GST tag followed by a PreScission protease cleavage site in *Escherichia coli* Rosetta (DE3) at 16 °C in LB medium. Cells were grown to OD<sub>600</sub> 0.8 and induced with 0.5 mM IPTG for 14 h, then harvested by centrifugation and resuspended with lysis buffer (20 mM HEPES pH 7.5, 2 M NaCl, 2 mM DTT and 2 mM EDTA). Cells were further disrupted by high pressure homogenizer and centrifuged at 18,000 *g* for 1 h. The soluble fraction containing fusion proteins was purified by GST-affinity chromatography (GE healthcare). After washed with washing buffer (20 mM HEPES pH 7.5, 2 M NaCl, 2 mM DTT), PreScission protease was supplemented into the column with a mass ratio of 1:100 and incubated at room temperature for 4 h to remove GST tag. The fractions containing target proteins were collected and diluted with buffer A (20 mM HEPES pH 7.5, 20 mM NaCl, 2 mM DTT), and loaded into Hitrap Heparin

column (GE healthcare), followed with gradient elution with buffer B (20 mM HEPES pH 7.5, 1 M NaCl, 2 mM DTT). The target fractions were collected and stored at -80 °C for further use.

### RNA transcription and purification

The respective RNAs were prepared by in vitro transcription and purified as previously reported<sup>51</sup>. Briefly, plasmids encoding an upstream T7 promoter and the respective RNAs were total-gene synthesized and confirmed by DNA sequencing (Wuxi Qinglan Biotechnology Inc, Wuxi, China). Using these plasmids as templates, plasmids encoding the respective RNA mutants or subdomains were generated by the Quick-change site-directed mutagenesis strategy and verified by DNA sequencing. The information of primers used was provided in Table S3. The double-stranded DNA fragment templates for in vitro RNA production were generated by PCR using an upstream forward primer targeted the plasmids and a downstream reverse primer specific to respective cDNAs. The RNAs were transcribed in vitro using T7 RNA polymerase and purified by preparative, non-denaturing polyacrylamide gel electrophoresis.

### Isothermal titration calorimetry (ITC)

ITC experiments were performed at 25 °C on a MicroCal PEAQ-ITC instrument (Malvern Instruments). Proteins and RNAs were buffer exchanged into 20 mM HEPES pH 7.5, 150 mM NaCl, 3 mM MgCl<sub>2</sub> with Superdex 75 size exclusion column (GE healthcare). The sample cell containing 300 µL of 30 µM RNA was titrated with 17 successive injections of 600 µM protein. Acquired titration curves were fitted with the Origin 7.0 program using the “one set of binding sites” binding model.

### flag-msi1 plasmid construction

The cDNA of human MS11 (NM\_002442.4) was amplified from the RNA of hNPCs with the primers: 5' GACGACGATGACAAGGGATCCATGGA GACTGACGCGCC 3' and 5' AAGCTTCTGCAGGTCGACTCAGTGGTACCATTGG 3', and cloned into pRK5-Flag vector by Gibson Assembly Cloning Kit (E5510S, NEB).

### Generation of mutant viruses

The substitutions construction and viral RNA transcription were performed as described above<sup>52</sup>. The primers used for site-directed mutagenesis for ZIKV- pMBSIM were: forward 5' ACCAAGCCCAAGTCCAGGCCG 3'; reverse 5'TTCCCAGCTTCTCCTGGG 3', for ZIKV-P2M were: forward 5'TAGTCAGGCCcgaAAgGCCATGGCAGC 3'; reverse 5'TGGGCTTGGTTCCAGC 3', for ZIKV-AGCU were: forward 5' TCA GGCCGAGctCGCCATGGCA 3'; reverse 5' CTATGGCTTGGTTTCCC 3' and for DENV4-DxM were: forward 5' CCgagaacGCCACGGTTT AGCAAAC 3'; reverse 5' CCTGACTaCAATAGCCTTTGGTGTTTGTTG 3'. The RNA was then transfected into BHK-21 cells using Lipofectamine 3000 reagent (Thermo Fisher Scientific). Culture supernatants were collected at 72 h post-transfection and used to infect C6/36 cells maintained in RPMI 1640 (Thermo Fisher Scientific) containing 2% FBS (Biowest). The C6/36 cells were cultured at 28 °C for 5–7 days and then the supernatants were dispensed into single-use aliquots, which were stored at -80 °C. Infectious virions were detected by plaque assay and viral antigen expression was detected by indirect immunofluorescent assay. The titers of virus stocks were then determined by plaque-forming assay, and the substitution sites were confirmed by RT-PCR and DNA sequencing.

### Plaque assay

Virus samples were serial-diluted by 10-fold with DMEM containing 2% FBS and 400 µl of each dilutes were added onto vero cells in 12-wells and incubated at 37 °C under 5% CO<sub>2</sub> for 2 h. The supernatants were then aspirated and 1 mL of DMEM containing 1% low melting point

agarose (Promega) and 2% FBS was added into each well. Four days post-infection, cells were fixed with 4% formaldehyde for 2 h at room temperature, followed by staining with 1% crystal violet solution for 30 min. After rinsing with water, the number of Plaques were counted for the calculations of virus titers.

### RNA pulldown

Mutated ZIKV 3'UTR DNA were synthesized and inserted into vector PUC57-KAN by Sangon (Shanghai, China) and then mutated ZIKV 3'UTR DNA fragment containing T7 promoters were amplified from them by PCR. The wild type ZIKV 3'UTR DNA fragment containing T7 promoter was amplified from the infectious clone. Then biotinylated WT and mutated ZIKV 3'UTR RNAs were synthesized by in vitro transcription using RiboMAX™ Large Scale RNA Production Systems-T7 (Promega) according to the manufacturer's instructions with the modification of the components of rNTPs were changed to: 8 mM GTP, 5 mM ATP, 5 mM CTP, 1.3 mM UTP, 0.7 mM Bio-11-UTP (Thermo Fisher Scientific). synthesized RNAs were purified using Purelink RNA mini kit (Thermo Fisher Scientific) and checked by agarose electrophoresis. BHK-21 cells were transfected with flag-*msi1*, forty-eight hours after transfection, cells were lysed by a lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 % NP-40, 5 mM EDTA, 10 % glycerol, 10 mM dithiothreitol, 1X cocktail protease inhibitor (YTHX Biotechnology) and 100 U/mL recombinant RNase inhibitor (Takara)). For the pulldown, 300ug of total protein from the cell lysates was precleared with 30uL Dynabeads MyOne Streptavidin T1 beads (Invitrogen). Biotinylated RNA was previously incubated at 60 °C for 5 min and then slowly cooled to room temperature. The precleared cell lysate was incubated with 5pmol RNA at 4 °C for 3 h. Subsequently, 30 uL Dynabeads MyOne Streptavidin T1 beads were added for another 2 h. After incubation, the beads were washed five times with lysis buffer. After washing, the beads were resuspended in 50 µL of 2X SDS loading buffer and boiled for 10 min at 98 °C. The samples were centrifuged and the supernatants were analyzed by Western Blot.

### RNA immunoprecipitation (RIP)

BHK-21 cells were transfected with flag-*msi1* plasmids. Twenty-four hours after transfection, cells were infected with ZIKV. Forty-eight hours after infection, cells were lysed by a lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 % NP-40, 5 mM EDTA, 10 % glycerol, 10 mM dithiothreitol, 1X cocktail protease inhibitor (YTHX Biotechnology) and 100 U/mL recombinant RNase inhibitor (Takara)). The cell lysates were subjected to immunoprecipitation with 0.5 ug anti-FLAG antibody or IgG control at 4 °C overnight and then mixed with 20 µL of gammabind G sepharose (GE Healthcare) and incubated for another 3 h. Beads were washed five times with lysis buffer and resuspended in 100 uL of lysis buffer. 50 uL of beads were used for Western Blot and 50 uL for RNA isolation using RNeasy Mini Kit (Qiagen). qRT-PCR was performed by use of the 3'UTR specific primers that amplify ZIKV 3'UTR (nucleotides 10623 to 10722 of ZIKV-Cambodia: forward 5'CCTGAAGTGGAGATCAGCTGTG 3'; reverse 5' GGTCTTCCACGG TCAATA 3').

### Replicon assay

The ZIKV replicon that carries the Renilla Luciferase gene was previously described<sup>53</sup>. The substitutions were constructed using the Q5 site-directed mutagenesis kit (NEB). Primers used for site-directed mutagenesis were as the same as primers used for the generation of mutant viruses described above. The replicon plasmids were linearized by restriction endonuclease digestion and purified by Phenol/Chloroform extraction. In vitro transcribed viral RNA was prepared using Ribomax T7 large scale RNA production kit (Promega) and purified using Purelink RNA mini kit (Thermo Fisher Scientific).  $2 \times 10^4$  cells

were seeded into each well of 48-well plates and incubated at 37 °C in 5% CO<sub>2</sub>. One day after seeding, 500 ng of each replicon RNA was transfected into each well of cells using Lipofectamine 3000 reagent (Thermo Fisher Scientific). The cell lysates were collected at the given time points and the luciferase activity assay was performed using the Renilla luciferase assay system (Promega) in a GloMax Discover system (Promega).

### Immunofluorescence analysis

Cells were fixed with 1% paraformaldehyde for 15 min and permeabilized with 0.2% Triton X-100 for 15 min, blocked with 3% bovine serum albumin for 2 h at 37 °C, and then incubated in the corresponding primary antibody (MSI1, 1:1000; sox2, 1:1000; ZIKV E, 1:1000), and then washed with PBS (3 × 5 min), followed by incubating in the secondary antibody (1:1000) at 37 °C for 1 h. After three washes with PBS, the cell nuclei were stained with DAPI (Merck D9542). Cells were photographed under an Olympus IX73 microscope. For the co-localization analysis, the samples were analyzed using a Nikon Eclipse Ti-U confocal microscope. Co-localization was calculated by Pearson's correlation coefficient (PCC) using Fiji with the Coloc2 plugin<sup>54</sup>.

### Western blot

The samples were fractionated by electrophoresis on 10% SDS-polyacrylamide gels, and resolved proteins were transferred onto PVDF membranes. After blocking with 5 % skimmed milk, the membranes were incubated with in the corresponding primary antibody (MSI1, 1:1500; FMRP, 1:1000; Actin, 1:10000; ZIKV E, 1:1000), and then washed with 0.05% tween-20 in PBS (4 × 5 min), followed by appropriate horseradish peroxidase-conjugated secondary antibodies (1:10000), and then washed with 0.05% tween-20 in PBS (4 × 5 min). Blots were developed using an enhanced chemiluminescence (ECL) kit (Thermo Fisher Scientific).

### Growth curve analysis

BHK-21 and SH-SY5Y cells were seeded onto 24-well plates and cultured overnight. For hNPC and U251, wells were previously coated with Matrigel Matrix (354248, corning) for 3 h at 37 °C with 5% CO<sub>2</sub>, the Matrigel Matrix then was discarded and cells were seeded. Cells were infected at indicted MOI. Infected cells were cultured at 37 °C with 5% CO<sub>2</sub> and culture supernatants were collected and the cells were fixed at the indicated time points. Viral RNA was extracted using Purelink RNA minikit (Thermo Fisher Scientific) and analyzed by qRT-PCR. qRT-PCR was performed using One Step PrimeScript™ RT-PCR kit (Takara) with the primers: forward 5' GGTCAGCGTCTCTCTAATAAACG3'; reverse 5' GCACCCTAGTGTCCACTTTTCC 3' and probe: FAM-AGCC ATGACCGACACCACACCGT-BQ1.

### RNA interference

A total of 30 pmol MSI1siRNA (sc-106836 santa cruz) or control siRNA (sc-37007 santa cruz) per well was transfected in a 24-well plate using RNAiMAX reagent (Thermo Fisher Scientific) as recommended by the manufacturer. The medium was changed at 6 h after transfection. Twenty-four hours after transfection, cells were infected with viruses at indicted MOI. Forty-eight hours after infection, cells supernatant and lysates were harvested for analysis.

### SAXS sample preparation, data collection, and processing

All the RNAs and proteins were buffer exchanged into SAXS buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 3 mM MgCl<sub>2</sub>, 5 mM DTT and 5% v/v glycerol) with size exclusion chromatography. To prepare RRM12-xrRNA2 complex, ZIKV xrRNA2 and MSI1 RRM12 were incubated at room temperature with a molar ratio at 1:3, and further purified through Superdex 75 size exclusion column with SAXS buffer. All the samples were concentrated to ~3 mg/ml for SAXS experiment.



Concentration series measurements (4- and 2-fold dilution and stock solution) were carried out and no aggregation or repulsion were observed. SAXS data was collected at room temperature at the beamline 12 ID-B of the Advanced Photon Source, Argonne National Laboratory as previously described<sup>51</sup>. Briefly, the setup was adjusted to achieve scattering  $q$  values of  $0.05 < q < 0.089 \text{ \AA}^{-1}$ , where  $q = (4\pi/\lambda) \sin\theta$ , and  $2\theta$  is the scattering angle. Thirty 2D images for each buffer or sample were recorded using a flow cell with the exposure time of 1 second. No radiation damage was observed as confirmed by the absence of systematic signal changes in sequentially collected X-ray scattering images. The 2D images were reduced to one-dimensional scattering profiles by MATLAB onsite. The scattering profile, the forward scattering intensity  $I(0)$  and the radius of gyration ( $R_g$ ), the pair distance distribution function (PDDF) as well as the maximum dimension,  $D_{\max}$  of the molecules were calculated using the same procedures as described before. The Volume-of-correlation ( $V_c$ ) was calculated by using the program Scatter, and the molecular weights of molecules were calculated on a relative scale using the  $R_g/V_c$  power law developed by Rambo et al.<sup>55</sup>. The theoretical scattering intensities of the atomic structure models were calculated and fitted to the experimental scattering profile using CRY SOL<sup>56</sup>.

### ab initio shape reconstruction and all-atom 3D atomic modeling

Low-resolution bead models of MSII RRM12, xrRNA2, and their complex were built up with the program DAMMIN, which generates models represented by an ensemble of densely packed beads, using SAXS scattering data within the  $q$  range of  $0.006\text{--}0.3 \text{ \AA}^{-1}$ . Forty independent runs were performed, and the resulting models were subjected to averaging by DAMAVER and were superimposed by SUPCOMB based on the normalized spatial discrepancy (NSD)<sup>38</sup>.

The atomic model of human MSII RRM12 was built up with MODELLER using the NMR structure of Mouse RRM1 (PDB ID: 2RS2) and RRM2 (PDB ID: 5X3Z) as templates<sup>16,19</sup>. The atomic model of MSII RRM12 was further refined against SAXS data with Xplor-NIH, which RRM1 (20–96) and RRM2 (109–191) were kept as rigid bodies to translate and rotate and the linker between (97–108) was allowed to translate or rotate freely. The atomic model of ZIKA xrRNA2 was built up with the FARFAR2, using the crystal structure of ZIKA xrRNA1 (PDB ID: STPY) along with NMR structure of AGAA tetraloop (PDB ID: 1K4A) as templates<sup>32,57</sup>.

The multiphase ab initio modeling program MONSA was used to obtain dummy atom model of RRM12-xrRNA2 complex. Twenty independent runs gave reproducible models, and the typical model is calculated using DAMAVER suite<sup>38</sup>.

### Hydrogen-deuterium exchange mass spectrometry (HDX-MS)

For HDX-MS experiment, 20  $\mu\text{L}$  of 5 mg/ml MSII RRM12 alone or in the presence of xrRNA2 were prepared. The pH of the buffer was 7.6, which contained 20 mM HEPES, 100 mM KCl, 3 mM  $\text{MgCl}_2$ , and 0.5 mM TCEP. To initiate deuterium labeling, 5  $\mu\text{L}$  of each 5 mg/ml protein solution was diluted with 45  $\mu\text{L}$  of labeling buffer (contents, 99%  $\text{D}_2\text{O}$ , pH 7.2) at room temperature for 60 s, 90 s, and 300 s. The reaction was quenched with 50  $\mu\text{L}$  of ice-cold quench buffer (1% (v/v) formic acid in water solution, 100%  $\text{H}_2\text{O}$ ). The reaction tube was then put on ice. 10  $\mu\text{L}$  of 1  $\mu\text{M}$  pepsin solution was added for digestion. After 5 min, the sample was centrifugated and placed into the auto-sampler of the Ultimate 3000 UPLC system (Thermo, CA, USA) for injection. The column and samples were precooled on ice. Fifty microlitres of sample was then loaded onto and separated by a ACQUITY UPLC 1.7  $\mu\text{M}$  BEHC18  $1.0 \times 50 \text{ mm}$  column (Waters). The peptides were eluted by a 16-min gradient of acetonitrile (1 to 50%) in 0.1% formic acid at 100  $\mu\text{L}/\text{min}$  with Thermo-Dionex Ultimate 3000 HPLC system. Mass spectrometry analysis was performed on Q Exactive Orbitrap mass spectrometer (Thermo, CA). The peptides were identified by using an in-house Proteome Discover (version PD1.4, Thermo Fisher Scientific), and HDX-

MS data were processed by HDEaminer from Thermo Fisher Scientific. The peptides were assigned as high or medium confidence by using HDEaminer (Sierra Analytics). The confidence level is calculated using a number of factors, including signal to noise and how well the theoretical isotope cluster matches the actual data<sup>58</sup>. The deuterium percentages were computed with fully deuterated samples. The fully deuterated samples were prepared by incubating sample with deuterium-exchange solution at room temperature for 24 h, and then quench the reaction with 36  $\mu\text{L}$  of optimal quench solution. The deuterium uptake was determined by monitoring shifts of the centroid peptide isotopic distribution by using the program HDEaminer (Thermo Fisher Scientific). The hydrogen/deuterium exchange difference of each peptide between protein alone and protein with ligand was manually checked. The experiments were repeated three times to ensure repeatability and deliver an estimate of the precision in the measurement.

### Integrative modeling of MSII RRM12-xrRNA2 complex

The structure of RRM12-xrRNA2 complex was predicted using the integrative modeling platform HADDOCK2.4 package<sup>39,40</sup>. The individual RRM1 (aa: 20-96) and RRM2 (aa: 109-190) domains rather than the integral RRM12 molecule were docked on xrRNA2, which allows large amplitude motions between RRM12 to occur during docking. The restraints used to guide sampling and evaluate the docked poses include: (1) The scattering profile and radius of gyration ( $R_g$ ) of the RRM12-xrRNA2 complex by SAXS; (2) The putative interaction interfaces between RRM12 and xrRNA2 derived from HDX-MS and mutagenesis data, and prior structural information on MSII RRM-ssRNA<sup>GUA</sup> complex (PDB ID: 2RS2 and 5X3Z). To facilitate extensive sampling in the conformational space of the RRM12-xrRNA2 complex, we increase the number of structures at the initial stage from 1000 (default value) to 10,000, and increase the steps of sampling at stages of docking (Table S6). From the xrRNA2 atomic model, the nucleobase of  $\text{G}^5$  within pMBS1 forms a canonical Watson-Crick base pair with  $\text{C}^{42}$  in PK1. The importance of PK1 to the fold and functional integrity of flaviviral xrRNAs have been described before<sup>59</sup>.  $\text{G}^5$  might be embedded in the core of xrRNA2 3D ring-like structure and solvent inaccessible to RRM12 (Fig. S10A-C), thus only the  $\text{U}^3\text{A}^4$  dinucleotides within pMBS1 are likely to interact with RRM12 during docking. The nucleotides 1–4 of xrRNA2 are allowed to be flexible during refinement.

Following the spirit of Ambiguous Interaction Restraints exploited by HADDOCK<sup>40</sup>, we defined the residues which are high solvent accessible and evidenced to be involved in binding as active residues and set the solvent-accessible neighbors of active residues as passive residues. The detailed information regarding the active and passive residues for each interface can be found in Table S7.

The SAXS scattering profile and the 3D shape envelope of MSII RRM12-xrRNA2 complex were used to score the representative models from the respective clusters generated by HADDOCK. The program CRY SOL was employed to calculate the theoretical scattering profile of the respective models and fit with the experimental scattering profile<sup>56</sup>.

### Structural illustration

The bead models generated by DAMMIN were transformed into volumetric maps using Situs<sup>60</sup>. All of the illustrations of atomic models were generated using the Chimera<sup>61</sup> or PyMOL Molecular Graphics System, version 1.3, Schrödinger, LLC.

### Statistical analysis

Sample size was estimated on the basis of similar research reported in the literature and no statistical method was used to predetermine sample size. No data were excluded from analyzes. All experiments were performed using at least 3 biological replicates to ensure

reproducibility. All samples were analyzed equally with no subsampling. Therefore, there was no requirement for randomization. Investigators were generally not blinded as the experimental conditions required investigators to know the identity of the samples. Description of the statistical tests performed for each experiment and the *p*-values can be found in the corresponding figure legends. Dots on the bar graphs represent the values from an individual replicate of the experiment. All data were analyzed using Prism software (GraphPad). For RNA enrichment, PCC, and virus infection rate data, unpaired two-sided Student's *t* test was performed to estimate the statistical significance between two groups. For virus growth curve analysis, data were analyzed by two-way ANOVA with Bonferroni correction for multiple comparisons. Representative data from repeated independent experiments are presented as mean  $\pm$  standard deviation (SD) with triplicate samples. Detailed statistical results can be found in the Source Data file.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All data are available with the article, Supplementary Information or Source Data file. Source data are provided as a Source Data file. Source data are provided with this paper.

### References

- Rubin, E. J., Greene, M. F. & Baden, L. R. Zika virus and microcephaly. *N. Engl. J. Med.* **374**, 984–985 (2016).
- Li, C. et al. Zika virus disrupts neural progenitor development and leads to microcephaly in mice. *Cell Stem Cell* **19**, 672 (2016).
- Li, H. et al. Zika virus infects neural progenitors in the adult mouse brain and alters proliferation. *Cell Stem Cell* **19**, 593–598 (2016).
- Tang, H. et al. Zika virus infects human cortical neural progenitors and attenuates their growth. *Cell Stem Cell* **18**, 587–590 (2016).
- Gabriel, E. et al. Recent Zika virus isolates induce premature differentiation of neural progenitors in human brain organoids. *Cell Stem Cell* **20**, 397–406.e395 (2017).
- Brault, J. B. et al. Comparative analysis between flaviviruses reveals specific neural stem cell tropism for Zika virus in the mouse developing neocortex. *EBioMedicine* **10**, 71–76 (2016).
- Muffat, J. et al. Human induced pluripotent stem cell-derived glial cells and neural progenitors display divergent responses to Zika and dengue infections. *Proc. Natl Acad. Sci. USA* **115**, 7117–7122 (2018).
- Sakakibara, S. et al. Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. *Dev Biol* **176**, 230–242 (1996).
- Good, P. et al. The human Musashi homolog 1 (MS1) gene encoding the homologue of Musashi/Nrp-1, a neural RNA-binding protein putatively expressed in CNS stem cells and neural progenitor cells. *Genomics* **52**, 382–384 (1998).
- Sakakibara, S. & Okano, H. Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. *J. Neurosci.* **17**, 8300–8312 (1997).
- Imai, T. et al. The neural RNA-binding protein Musashi1 translationally regulates mammalian *numb* gene expression by interacting with its mRNA. *Mol. Cell Biol.* **21**, 3888–3900 (2001).
- Nakamura, M., Okano, H., Blendy, J. A. & Montell, C. Musashi, a neural RNA-binding protein required for *Drosophila* adult external sensory organ development. *Neuron* **13**, 67–81 (1994).
- Okabe, M., Imai, T., Kurusu, M., Hiromi, Y. & Okano, H. Translational repression determines a neuronal potential in *Drosophila* asymmetric cell division. *Nature* **411**, 94–98 (2001).
- Okano, H., Imai, T. & Okabe, M. Musashi: a translational regulator of cell fate. *J. Cell Sci.* **115**, 1355–1359 (2002).
- Sakakibara, S. et al. RNA-binding protein Musashi family: roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. *Proc. Natl Acad. Sci. USA* **99**, 15194–15199 (2002).
- Iwaoka, R. et al. Structural insight into the recognition of r(uag) by musashi-1 rbd2, and construction of a model of musashi-1 rbd1-2 bound to the minimum target RNA. *Molecules* **22**, <https://doi.org/10.3390/molecules22071207> (2017).
- Uren, P. J. et al. RNA-binding protein musashi1 is a central regulator of adhesion pathways in glioblastoma. *Mol. Cell Biol.* **35**, 2965–2978 (2015).
- Zearfoss, N. R. et al. A conserved three-nucleotide core motif defines Musashi RNA binding specificity. *J Biol Chem* **289**, 35530–35541 (2014).
- Ohyama, T. et al. Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. *Nucleic Acids Res.* **40**, 3218–3231 (2012).
- Battelli, C., Nikopoulos, G. N., Mitchell, J. G. & Verdi, J. M. The RNA-binding protein Musashi-1 regulates neural development through the translational repression of p21WAF-1. *Mol. Cell Neurosci.* **31**, 85–96 (2006).
- Horisawa, K., Imai, T., Okano, H. & Yanagawa, H. 3'-Untranslated region of doublecortin mRNA is a binding target of the Musashi1 RNA-binding protein. *FEBS Lett.* **583**, 2429–2434 (2009).
- de Sousa Abreu, R. et al. Genomic analyses of musashi1 downstream targets show a strong association with cancer-related processes. *J. Biol. Chem.* **284**, 12125–12135 (2009).
- Zhu, Z. et al. Comparative genomic analysis of pre-epidemic and epidemic Zika virus strains for virological factors potentially associated with the rapidly expanding epidemic. *Emerg. Microbes Infect.* **5**, e22 (2016).
- Liu, Z. Y. et al. Viral RNA switch mediates the dynamic control of flavivirus replicase recruitment by genome cyclization. *Life* **5**, <https://doi.org/10.7554/eLife.17636> (2016).
- Shan, C. et al. A live-attenuated Zika virus vaccine candidate induces sterilizing immunity in mouse models. *Nat. Med.* **23**, 763–767 (2017).
- Chavali, P. L. et al. Neurodevelopmental protein Musashi-1 interacts with the Zika genome and promotes viral replication. *Science* **357**, 83–88 (2017).
- Klase, Z. A. et al. Zika fetal neuropathogenesis: etiology of a viral syndrome. *PLoS Negl. Trop. Dis.* **10**, e0004877 (2016).
- Schneider, Ad. B. et al. Molecular evolution of Zika virus as it crossed the Pacific to the Americas. *Cladistic* **33**, 1–20 (2017).
- Schneider, A. B. & Wolfinger, M. T. Musashi binding elements in Zika and related Flavivirus 3'UTRs: a comparative study in silico. *Sci. Rep.* **9**, 6911 (2019).
- Soto-Acosta, R. et al. Fragile X mental retardation protein is a Zika virus restriction factor that is antagonized by subgenomic flaviviral RNA. *Life* **7**, <https://doi.org/10.7554/eLife.39023> (2018).
- Shan, C. et al. An infectious cDNA clone of Zika virus to study viral virulence, mosquito transmission, and antiviral inhibitors. *Cell Host Microbe* **19**, 891–900 (2016).
- Akiyama, B. M. et al. Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science* **354**, 1148–1152 (2016).
- Wu, H., Henras, A., Chanfreau, G. & Feigon, J. Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proceedings of the National Academy of Sciences* **101**, 8307–8312 (2004).
- Bottaro, S. & Lindorff-Larsen, K. Mapping the universe of RNA tetraloop folds. *Biophys. J.* **113**, 257–267 (2017).

35. Daubner, G. M., Clery, A. & Allain, F. H. RRM-RNA recognition: NMR or crystallography...and new findings. *Curr. Opin. Struct. Biol.* **23**, 100–108 (2013).
36. Afroz, T., Cienikova, Z., Clery, A. & Allain, F. H. T. One, two, three, four! how multiple rrms read the genome sequence. *Methods Enzymol.* **558**, 235–278 (2015).
37. Dimitrova-Paternoga, L., Jagtap, P. K. A., Chen, P. C. & Hennig, J. Integrative structural biology of protein-RNA complexes. *Structure* **28**, 6–28 (2020).
38. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879–2886 (1999).
39. van Zundert, G. C. P. et al. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
40. Dominguez, C., Boelens, R. & Bonvin, A. M. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737 (2003).
41. Chapman, E. G. et al. The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production. *Science* **344**, 307–310 (2014).
42. Filomatori, C. V. et al. Dengue virus genomic variation associated with mosquito adaptation defines the pattern of viral non-coding RNAs and fitness in human cells. *PLoS Pathog.* **13**, e1006265 (2017).
43. Manokaran, G. et al. Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science* **350**, 217–221 (2015).
44. Schnettler, E. et al. Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells. *J Virol* **86**, 13486–13500 (2012).
45. Schnettler, E. et al. Induction and suppression of tick cell antiviral RNAi responses by tick-borne flaviviruses. *Nucleic Acids Res.* **42**, 9436–9446 (2014).
46. Slonchak, A. et al. Zika virus noncoding RNA suppresses apoptosis and is required for virus transmission by mosquitoes. *Nat. Commun.* **11**, 2205 (2020).
47. Berman, H. M. et al. The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
48. Cléry, A., Blatter, M. & Allain, F. H. T. RNA recognition motifs: boring? not quite. *Curr. Opin. Struct. Biol.* **18**, 290–298 (2008).
49. Kinney, R. M. et al. Avian virulence and thermostable replication of the North American strain of West Nile virus. *J. Gen. Virol.* **87**, 3611–3622 (2006).
50. Deng, Y. Q. et al. A broadly flavivirus cross-neutralizing monoclonal antibody that recognizes a novel epitope within the fusion loop of E protein. *PLoS ONE* **6**, e16059 (2011).
51. Zhang, Y. et al. Long non-coding subgenomic flavivirus RNAs have extended 3D structures and are flexible in solution. *EMBO Rep.* **20**, e47016 (2019).
52. Yuan, L. et al. A single mutation in the prM protein of Zika virus contributes to fetal microcephaly. *Science* **358**, 933–936 (2017).
53. Li, C. et al. 25-hydroxycholesterol protects host against Zika virus infection and its associated microcephaly in a mouse model. *Immunity* **46**, 446–456 (2017).
54. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
55. Rambo, R. P. & Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481 (2013).
56. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).
57. Wu, H. et al. A novel family of RNA tetraloop structure forms the recognition site for *Saccharomyces cerevisiae* RNase III. *EMBO J.* **20**, 7240–7249 (2001).
58. Masson, G. R. et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods* **16**, 595–602 (2019).
59. Niu, X. et al. Pseudoknot length modulates the folding, conformational dynamics, and robustness of Xrn1 resistance of flaviviral xrRNAs. *Nat. Commun.* **12**, 6417 (2021).
60. Wriggers, W. Conventions and workflows for using Situs. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 344–351 (2012).
61. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

## Acknowledgements

We thank Dr. Andrew Davidson from the University of Bristol for kind suggestion and discussion, Drs. Xiaolin Tian and Haiteng Deng in Center of Protein Analysis Technology, Tsinghua University for HDX-MS analysis, Cuiyan Zhou in Protein Preparation and Identification Facility, Tsinghua University for assistance with ITC experiments, and the staff at the beamline 12-ID-B, Argonne National Laboratory, USA, and beamline BL19U2, Shanghai Synchrotron Radiation Facility, China, for assistance during SAXS data collection. This work was supported in part by the National Key Research and Development Project of China (2018YFA0900801 (H.Z.), 2021YFC2302400 (C.-F.Q.), and 2021YFA1301500 (X.F.)), the National Natural Science Foundation (NSFC) (32100127 (R.L.), U1832215 (X.F.), and 31770190 (C.-F.Q.)), the Postdoctoral Science Foundation of China (2019M664017) (X.C.), C.-F.Q. was supported by the National Science Fund for Distinguished Young Scholars (81925025), the Innovative Research Group (81621005) from the NSFC, and the Innovation Fund for Medical Sciences (2019-I2M-5-049) from the Chinese Academy of Medical Sciences.

## Author contributions

C.F.Q. and X.F. conceived, designed and supervised the study. X.C., Y.W. and Z.X. performed the majority of the experiments and analyzed the data. X.Z. performed the SAXS measurements. Q.M. and Z.W. conducted immunostaining. M.C. and R.L. performed the sequence alignment and analysis under the guidance of H.Z. X.L. edited the draft. All authors read and approved the contents of the manuscript.

## Competing interests

C.F.Q. and X.F. have filed a patent related to the major findings reported in this study. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-36838-w>.

**Correspondence** and requests for materials should be addressed to Xianyang Fang or Cheng-Feng Qin.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2023