Cryo-EM structures of Lassa and Machupo virus polymerases complexed with cognate regulatory Z proteins identify targets for antivirals

Xin Xu^{1,2,11}, Ruchao Peng^{1,2,11}, Qi Peng^{2,11}, Min Wang², Ying Xu³, Sheng Liu², Xiaolin Tian⁴, Haiteng Deng⁴, Yimin Tong⁵, Xiaoyou Hu^{1,5}, Jin Zhong^{1,5,6}, Peiyi Wang⁷, Jianxun Qi², George F. Gao^{1,2,8} and Yi Shi^{1,2,8,9,10} ✉

Zoonotic arenaviruses can lead to life-threatening diseases in humans. These viruses encode a large (L) polymerase that transcribes and replicates the viral genome. At the late stage of replication, the multifunctional Z protein interacts with the L polymerase to shut down RNA synthesis and initiate virion assembly. However, the mechanism by which the Z protein regulates the activity of L polymerase is unclear. Here, we used cryo-electron microscopy to resolve the structures of both Lassa and Machupo virus L polymerases in complex with their cognate Z proteins, and viral RNA, to 3.1–3.9 Å resolutions. These structures reveal that Z protein binding induces conformational changes in two catalytic motifs of the L polymerase, and restrains their conformational dynamics to inhibit RNA synthesis, which is supported by hydrogen–deuterium exchange mass spectrometry analysis. Importantly, we show, by in vitro polymerase reactions, that Z proteins of Lassa and Machupo viruses can cross-inhibit their L polymerases, albeit with decreased inhibition efficiencies. This cross-reactivity results from a highly conserved determinant motif at the contacting interface, but is affected by other variable auxiliary motifs due to the divergent evolution of Old World and New World arenaviruses. These findings could provide promising targets for developing broad-spectrum antiviral drugs.

Arenaviruses pose great public concerns for human health, especially in many developing and undeveloped countries without advanced hygiene facilities^{1–3}. As a representative species, Lassa virus (LASV) infection usually causes flu-like symptoms. Patients who progress to severe cases would then develop acute haemorrhagic fever and neurological disorders, which may lead to multiple-organ failure and death⁴. LASV is endemic in many African countries, resulting in 100,000–300,000 human infection cases annually, including approximately 5,000 deaths^{4,5}. Currently, there are no specific drugs or vaccines available for most arenaviruses, except for a Junin virus vaccine⁶. The broad-spectrum antiviral drug ribavirin has shown efficacy against LASV only if administered at the very early stage of infection⁷. Therefore, it is urgent to develop more efficient vaccines and antiviral drugs for arenaviruses.

Arenaviruses are enveloped RNA viruses comprising the family *Arenaviridae*, which includes four genera: *Mammarenavirus*, *Reptarenavirus*, *Hartmanivirus* and *Antennavirus*⁸. All human-infecting viral species are included in the genus *Mammarenavirus*. Typically, the genome of arenaviruses consists of two ambisense RNA segments and encodes four viral proteins: the envelope glycoprotein precursor; nucleoprotein; large (L) RNA-dependent RNA polymerase (RdRp); and zinc finger (Z) matrix protein⁸. Both transcription and replication are mediated by the L polymerase in the context of the ribonucleoprotein complex (RNP)^{9,10}. Each viral RNA (vRNA) segment contains two open reading frames arranged in

opposite directions. The expression of early and late genes could thus be regulated by different coding orientations¹⁰. The L protein and nucleoprotein are expressed immediately after virus entry to support genome replication and protein production. At the later stage, the multifunctional Z protein would be produced to counteract the host immunity¹¹ and to slow down RNA synthesis by L polymerase^{9,12}, which facilitates the balance of the infection process and initiates virion assembly¹³. In addition, the glycoprotein precursor would also be expressed to enable the complete assembly and egress of progeny virions⁸. Previous studies have shown that the Z protein could inhibit L polymerase activity both in vitro and in vivo^{12,14,15}, but the underlying mechanism is unclear.

In this work, we determined the structures of LASV and Machupo virus (MACV) L polymerases in complex with their cognate Z proteins, as well as the vRNA promoter, by cryo-electron microscopy (cryo-EM) reconstruction. These structures, combined with biochemical evidence, reveal both conserved and unique features underpinning the L–Z interactions for Old World and New World arenavirus groups, allowing mechanistic insights into the regulation of arenavirus replication and suggesting a promising strategy for antiviral drug design.

Results

Overall structures of arenavirus L–Z/L–Z–vRNA complexes. The structures of LASV and MACV L–Z complexes were determined at

¹Savai Medical School, University of Chinese Academy of Sciences, Beijing, China. ²CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. ³College of Life Science, Nanjing Agricultural University, Nanjing, China. ⁴MOE Key Laboratory of Bioinformatics, School of Life Sciences, Tsinghua University, Beijing, China. ⁵CAS Key Laboratory of Molecular Virology and Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences, Shanghai, China. ⁶ShanghaiTech University, Shanghai, China. ⁷Department of Biology, Southern University of Science and Technology, Shenzhen, China. ⁸Center for Influenza Research and Early-Warning (CASCIRE), CAS-TWAS Center of Excellence for Emerging Infectious Diseases (CEEID), Chinese Academy of Sciences, Beijing, China. ⁹Chongqing Key Laboratory of Neurodegenerative Diseases, Chongqing General Hospital, University of Chinese Academy of Sciences, Chongqing, China. ¹⁰College of Basic Medicine, Jilin University, Changchun, China. ¹¹These authors contributed equally: Xin Xu, Ruchao Peng, Qi Peng. ✉e-mail: shiyi@im.ac.cn

3.9 and 3.4 Å overall resolutions, respectively (Extended Data Figs. 1 and 2 and Supplementary Table 1). In each complex, one copy of the Z protein binds to an L polymerase at the interface between RdRp and polymerase acidic subunit carboxy (C)-terminal-like (PA-C-like) regions (Fig. 1a–c). Similar to the apo polymerase structures, the RdRp core is highly stable, whereas the peripheral domains display different extents of flexibility¹⁶. The C-terminal domain (CTD) could not be resolved in both structures of the L–Z heterodimeric complex, which contains the putative cap-binding domain that is critical for transcription¹⁷ (Fig. 1a–c). In the cryo-EM images of the MACV L–Z complex, a small fraction of dimeric complex, (L–Z)₂, was observed, in which a homodimer of L polymerase bound to two copies of Z protein (Extended Data Fig. 2c,i). This dimeric configuration of L protein stabilized the CTD to facilitate its structural visualization. The endonuclease domain was resolved at moderate resolution for both LASV and MACV L proteins, with obvious conformational heterogeneity (Fig. 1b,c), consistent with its functionality for cap snatching¹⁸. The entire arenavirus-specific pendant insertion domain was missing in the density map of the LASV L–Z complex, whereas it was well resolved in the MACV L–Z complex (Extended Data Fig. 3), suggesting different conformational states of the two structures.

The Z protein consists of a stable central RING domain and two flexible arms in the amino (N) and C termini¹⁹. In the structures of L–Z complexes, only the RING domain and small portions of the N- and C-terminal loops were observed, denoted as the L-binding domain (LBD), which mainly interacted with the palm domain of RdRp (Fig. 1 and Extended Data Fig. 3b). Previous electrophoretic mobility shift assay revealed that binding of the Z protein did not prevent vRNA recruitment by L polymerase¹⁴. To verify this observation, we also determined the structure of MACV L bound to Z protein and the vRNA promoter (Fig. 1d, Extended Data Figs. 4 and 5 and Supplementary Table 2). The MACV L–Z–vRNA complex exists as both monomers (L–Z–vRNA) and dimers (L–Z–vRNA)₂. In both complex forms, clear density for the 3' vRNA was observed in a groove formed by the PA-C-like region and the thumb domain of RdRp, consistent with the structure of the L–vRNA complex without Z protein binding¹⁶ (Fig. 1e and Supplementary Fig. 1). In the presence of 3' vRNA, the density for the α -ribbon helices and pendant insertion domain became disordered, suggesting conformational changes of this region (Extended Data Fig. 3c and Supplementary Fig. 1). Also, the binding of 3' vRNA substantially promoted dimerization of MACV L polymerase, which allowed us to reconstruct the CTD to high resolution (Extended Data Figs. 4 and 5). The main chain trace of the cap-binding domain was well defined in the density map, whereas the N- and C-terminal arms of the Z protein still could not be visualized, suggesting that the two terminal arms do not contribute to interactions with L polymerase (Fig. 1f). This is consistent with the previous observation that the central RING domain, but not the terminal loops, is essential to maintain the inhibitory activity of Z protein for L polymerase in replicon cells²⁰.

To analyse the potential effect of Z protein binding on RNA synthesis, we modelled the template and product RNA strands into the catalytic centre of the MACV L protein (Extended Data Fig. 6), based on the structure of influenza polymerase in the elongation state (Protein Data Bank (PDB) 6QCT)²¹. According to this model, the product–template duplex is separated by the polymerase basic subunit N-terminal-like (PB2-N-like) region, and the nascent RNA strand emerges out of the catalytic chamber through the large cleft between the finger and thumb domains of RdRp. The Z protein binds to the distal end of the palm domain (Extended Data Fig. 6a), which is far from the entry/exit tunnels for RNA strands and nucleoside 5'-triphosphate (NTP) substrates. This implies that Z protein does not sterically block RNA elongation or NTP substrate feeding, but may inactivate the polymerase by allosteric effects instead. This

hypothesis is supported by the observation that Z protein universally inhibited the production of both fully extended and prematurely terminated RNA products (Supplementary Fig. 2).

Molecular determinants for L–Z interaction and inhibition.

The Z protein binding site on L polymerase involves RdRp and the PA-C- and PB2-N-like regions, among which the palm domain of RdRp and the core lobe of the PA-C-like region contribute most of the interactions (Fig. 2a). The LBD of the Z protein consists of a central helix and a few terminal loops stabilized by two conserved zinc fingers. The central helix is not substantially involved in interactions with L proteins (Fig. 2b). Instead, three short loops of Z protein protrude into the grooves of L polymerase at the interface, resulting in three discrete contacting sites (Fig. 2c–i).

At contacting site 1, two aromatic residues of the Z protein (Trp35 and Phe36 for LASV and Trp43 and Phe44 for MACV) insert into a hydrophobic cavity formed by the RdRp and PA-C-like regions of L polymerase (Fig. 2d and Extended Data Fig. 7a,b). Residue Trp35 of LASV Z contributes most of the Van der Waals contacts with the surrounding non-polar residues in L protein and is strictly conserved for all mammarenaviruses (Extended Data Fig. 7 and Supplementary Table 3). Apart from the non-polar interactions, residue Phe36 is also involved in π -cation interactions with Lys263 in the PA-C-like region. Moreover, the carbonyl oxygen of Ser33 forms a hydrogen bond with the main chain of Met649 to further stabilize the interaction (Fig. 2d). The corresponding site in the MACV L–Z complex displays a similar interacting profile but with fewer atomic contacts (Fig. 2g and Supplementary Tables 3 and 4). These observations suggest a more stable hydrophobic interface in this region for the LASV L–Z complex than for the MACV L–Z complex, which may explain the higher affinity (around ten-fold difference) of LASV L–Z interaction, as revealed by bio-layer interferometry (BLI) assays (Fig. 3a,b). Alanine substitution of the highly conserved Trp residue abolished L–Z interactions for both LASV and MACV and resulted in loss of the inhibitory effect (Fig. 3 and Supplementary Fig. 3), consistent with the previous observations in LCMV and Tacaribe virus (TCRV) replicon cells^{20,22}. The p.Phe44Ala substitution of MACV Z also completely impaired its binding and inhibition capacity for L polymerase, while the p.Phe36Ala mutant of LASV Z retained efficient binding to LASV L, albeit with slightly reduced affinity (around fourfold difference) (Fig. 3a). This difference may result from the less stable hydrophobic interface in the MACV L–Z complex, which thus requires the presence of both residues Trp and Phe to stabilize the interactions. Despite the good affinity to the L protein, the p.Phe36Ala mutant of LASV Z exhibited highly compromised inhibitory efficiency (<30% of the inhibitory efficiency of the wild-type Z) for both transcription and replication by LASV L using a 19-nucleotide 3' vRNA template (Fig. 3d,e). This evidence further implies allosteric modulation on L polymerase by the Z protein, which directly inactivates catalysis.

Since p.Phe36Ala substitution of LASV Z did not impair its binding to L polymerase, we further determined the structure of LASV L bound to 3' vRNA and the p.Phe36Ala Z protein mutant at 3.4 Å resolution (Extended Data Fig. 8). This structure better resolved the atomic details at the L–Z contacting interface and displayed highly similar interaction profiles to the complex with the wild-type Z protein. Further structural analysis on LASV L–Z interactions at sites 2 and 3 was mainly based on this structure. At these two sites, residue Phe30 of the Z protein interacts with Arg1380 in the RdRp through π -cation interactions, and residue Lys68 potentially forms a salt bridge with Asp1724 in the PB2-N-like region (Fig. 2e,f). A few Van der Waals contacts are also involved in both sites (Supplementary Tables 3 and 5). In the corresponding sites of the MACV L–Z complex, residue Arg36 of the Z protein engages two residues in the RdRp (that is, Thr1377 and Phe1378) via hydrogen bonds and π -cation interactions, respectively (Fig. 2h). The main

Fig. 1 | Overall structures of LASV and MACV L-Z complexes with and without vRNA. **a**, Schematic diagram of the domain structures for the L and Z proteins. Each domain is represented with a unique colour. The CTD of L polymerase is not resolved in the monomeric L-Z complex structures and is only visualized in the dimeric forms of the MACV L-Z and L-Z-vRNA complexes. Most of the flexible N- and C-terminal arms of Z proteins are not observed either, and are represented by waved lines. The Z protein interacts with three parts of L polymerase, which are connected by dashed lines. Endo, endonuclease domain. **b,c**, Cryo-EM density maps of LASV (**b**) and MACV (**c**) L-Z monomeric complexes. The structures are coloured by domain, with the same colour scheme as in **a**. **d**, Composite map of the dimeric complex of MACV L bound to the Z protein and vRNA. This complex contains a homodimer of L polymerase bound with two copies of 3' vRNA and two Z proteins, (L-Z-vRNA)₂. One of the L-Z-vRNA protomers is coloured by domain, whereas the other is coloured in light colours to facilitate visualization of the dimeric assembly. **e**, Magnified view of the 3' vRNA binding site indicated by a blue dashed box in **d**. The 5' portion of the 3' vRNA strand could not be resolved and is represented by a black dashed line. **f**, Magnified view of the Z protein binding site in the MACV L-Z-vRNA dimeric complex indicated by a red dashed box in **d**. The cap-binding domain (orange) is well resolved in this complex, whereas the N- and C-terminal arms of the Z protein still could not be visualized (as indicated by dashed lines).

chain of His73 hydrogen bonds to the side chain of Arg691 in the PA-C-like region, and residue Trp76 forms a few Van der Waals contacts with Asn1712 and Phe1715 in the PB2-N-like region (Fig. 2i). These two sites reveal remarkable differences between Old World and New World mammarenaviruses but are quite conserved within the two individual groups, indicating the group-specific mechanisms utilized by Old World and New World mammarenaviruses

in these regions (Extended Data Fig. 7c). Substitution of residues at these two sites showed only slight or no effects on the binding affinity and inhibitory efficiency of LASV Z protein, whereas the p.Arg36Ala mutant of the MACV Z protein almost completely lost the capacity for binding and inhibiting MACV L polymerase (Fig. 3). These data together demonstrate the different molecular determinants of LASV and MACV for L-Z interaction and

Fig. 2 | L-Z protein interactions for LASV and MACV. **a**, Overall structure of MACV L polymerase bound to the Z protein. The density map was low-pass filtered to better reveal the shape of each structural part. The density of the Z protein (coloured in salmon) is set as transparent to enable visualization of the atomic model. **b**, Binding interface between L and Z proteins. The L polymerase is shown as the surface model and coloured by domain. The Z protein is shown as a cartoon model in which the two zinc-finger motifs (Zn 1 and 2) are represented by sticks and spheres. **c**, Magnified view of the L-Z contacting interface. The three contacting Z protein loops are highlighted, comprising three discrete interacting sites. **d-i**, Atomic interactions between LASV (**d-f**) and MACV (**g-i**) L polymerases and Z proteins for sites 1 (**d** and **g**), 2 (**e** and **h**) and 3 (**f** and **i**), as indicated by the coloured dashed boxes in **c**. The key interacting residues are shown as sticks and coloured by element. Hydrogen bonds and salt bridges are represented by black dashed lines, whereas π -cation interactions are represented by red dashed lines. The distances between contacting atoms (in Ångström) are labelled along with the dashed lines. Panel **d** was prepared based on the structure of the LASV L-Z complex (3.9 Å resolution). Panels **e** and **f** are based on the structure of LASV L bound to vRNA and the p.Phe36Ala Z protein mutant (3.4 Å resolution), which better resolves these two sites.

inhibition, suggesting the divergence of Old World and New World mammarenaviruses in evolution.

Conformational modulation of L polymerase by Z protein. A previous study indicated that Z protein binding may involve the catalytic motifs of L polymerase²³. We thus compared the structures of L polymerases in apo and Z protein-bound states to analyse the potential conformational changes within the catalytic centre. As a typical viral RdRp, the active site of arenavirus polymerase contains eight conserved motifs (motifs A–H) (Fig. 4a,b). None of them displays discernible conformational changes within the catalytic centre upon Z protein binding. Importantly, the distal ends of two catalytic motifs, D and E, are involved in interactions with the Z protein outside the catalytic centre (Fig. 4a,b). Motif E is located at contacting site 1, which accommodates the highly conserved Trp and Phe residues of the Z protein. Two hydrophobic residues in this motif undergo substantial conformational changes to reorganize a stable hydrophobic cavity that allows space for embracing the strictly conserved Trp (Fig. 4c,d). An adjacent loop in the PA-C-like region (motif G) is also reoriented to avoid a steric clash with the well-conserved Phe residue (Fig. 4d).

To understand how L polymerase mediates catalysis, we modelled the template/product RNA strands and the NTP substrate into the active site of the LASV L protein based on the structure of influenza virus polymerase in the pre-catalytic conformation (PDB 6SZV)²¹. According to this model, residue Ser1387 (motif E, numbered by LASV L) stabilizes the 3' end of the primer via hydrogen binding to the phosphate backbone. Two positively charged residues, Lys1375 (motif D) and Lys1195 (motif A) may interact with the γ -phosphate of the NTP substrate to support catalysis (Fig. 5a,b). These three residues are highly conserved for all segmented negative-sense RNA viruses and are also observed in the polymerase of coronaviruses—a group of positive-sense RNA viruses²⁴. Since the Z protein engages the distal end of motifs D and E, we hypothesized that these two motifs might be immobilized to restrict the conformational changes required for catalysis (Fig. 5c). To test this hypothesis, we performed hydrogen–deuterium exchange mass spectrometry (HDX-MS) assay for LASV L polymerase in the presence or absence of the Z protein. The entire motif D loop and part of motif E with its adjacent loop region could be detected by mass spectrometry, allowing us to analyse the effects of Z protein binding on the conformational dynamics of these catalytic motifs (Fig. 5c,d).

Fig. 3 | Mutagenesis studies of Z proteins and the effects on binding and inhibiting L polymerases. **a,b**, BLI binding kinetics of LASV (**a**) and MACV (**b**) wild-type and mutant Z proteins to cognate L polymerases. The Z proteins or mutants were immobilized, and serially-diluted L proteins were applied to test the binding. The data are representative of three independent experiments using different protein preparations. **c,f**, Titration of LASV (**c**) and MACV (**f**) wild-type Z proteins for inhibiting transcription by cognate L polymerases. Each data point indicates the mean value of three independent experiments using different protein preparations. The error bars represent s.d. The L polymerases were used at a working concentration of 0.4μ M. **d,g**, Inhibitory activities of wild-type and mutant Z proteins for replication and transcription by L polymerases. In these experiments, de novo replication and cap-dependent transcription assays were performed to test the inhibitory activities of the Z proteins. The wild-type and mutant Z proteins of LASV (**d**) and MACV (**g**) were used at $1.5\times$ and $1.7\times$ the IC_{50} concentrations, respectively, at which the wild-type proteins displayed $>90\%$ inhibition for polymerase activities. The bands of expected replication and transcription products are indicated by blue and red arrowheads, respectively. nt, nucleotides. **e,h**, Quantification of the inhibitory efficiencies on the polymerase activities shown in **d** and **g**, respectively. The data represent mean values (histograms) $\pm$ s.d. (error bars) of three independent experiments using different protein preparations. Statistical significance was determined by one-way analysis of variance ($***P < 0.001$). The exact P values are summarized in Supplementary Table 7.

Both peptide fragments displayed obviously reduced efficiencies for deuterium uptake in the presence of wild-type Z protein compared with L polymerase alone, suggesting the restricted mobility of these motifs upon Z protein binding (Fig. 5e). As a negative control, presence of the p.Trp35Ala Z protein mutant did not lead to obvious changes in the hydrogen–deuterium exchange kinetics of LASV L. In contrast, binding of the p.Phe36Ala mutant obviously reduced the

efficiency of deuterium uptake by the L polymerase—to a lesser extent compared with the wild-type Z protein (Fig. 5e). This observation indicates that the p.Phe36Ala mutant of LASV Z is capable of interfering with the conformational dynamics of L polymerase, which may retain substantial inhibitory activity for RNA synthesis. The poor inhibitory efficiency for transcription and replication using a 19-nucleotide template may arise from the short span

Fig. 4 | Z protein binding-induced conformational changes of L polymerases. a,b, LASV (**a**) and MACV (**b**) L polymerase catalytic motifs at the Z-contacting interface. The L and Z proteins are shown as cartoon and surface models, respectively. The eight conserved catalytic motifs (A–H) of RdRp are represented with different colours. The key catalytic residues and Z-interacting residues are shown as sticks and the catalytic metal ions are shown as spheres. **c,d**, Conformational changes of LASV (**c**) and MACV (**d**) L polymerase before and after Z protein binding. The structures of apo and Z protein-bound L polymerases are superimposed and magnified views of the L–Z interacting interface are shown. The apo polymerases are coloured in grey and the Z protein-bound forms are coloured by motif, with the same colour scheme as in **a** and **b**. The key interacting residues are shown as sticks and the conformational changes are indicated by arrows.

of the elongation process, which may not be able to resolve the relatively weaker effect compared with the wild-type Z protein (Fig. 3d,e). To verify this hypothesis, we re-conducted the *in vitro* RNA synthesis assays with a longer template (41 nucleotides), in which the p.Phe36Ala mutant exhibited a more potent inhibitory efficacy for both transcription and replication (~60% of the inhibitory efficiency of the wild-type Z protein) (Extended Data Fig. 9). Therefore, these data together demonstrate that the Z protein can allosterically regulate the conformational dynamics of L polymerase, which, as a result, inactivates its catalytic activity.

Cross-inhibition of LASV/MACV Z proteins to L polymerases.

Since the dominant contacting site (site 1) is conserved across all mammarenaviruses and the critical residues Trp and Phe are highly conserved, we hypothesized that Z proteins might be able to contact and inhibit the heterologous L polymerases. As expected, both LASV and MACV Z proteins could mutually interact with the other L polymerase and led to cross-inhibition for RNA synthesis, albeit with different kinetic features and inhibitory efficiencies (Fig. 6). The binding of MACV Z to LASV L ($K_d=53.4$ nM) displayed a higher dissociation rate than that for LASV L–Z

interactions ($K_d=0.99$ nM), resulting in approximately 50-fold reduction in the binding affinity (Figs. 3a and 6a and Supplementary Table 6). Accordingly, MACV Z is less efficient for inhibiting LASV L compared with its cognate Z protein, with a half-maximum inhibitory concentration (IC_{50}) that is roughly 3.5 times higher (Figs. 3c and 6c and Supplementary Fig. 2). Despite having a slightly higher affinity, LASV Z displayed a lower efficiency ($IC_{50}=8.9$ μ M) for inhibiting the activity of MACV L than MACV Z ($IC_{50}=2.4$ μ M) (Figs. 3f and 6b,d). Both Z proteins revealed slower association rates for binding heterologous L than those for interactions with cognate polymerases, suggesting suboptimal conformational adaptation in heterologous interactions (Supplementary Table 6). Given that the sequence identity between LASV and MACV L polymerases is only ~35%, the suboptimal cross-inhibition further highlights the certain specificity of L–Z interactions for different arenavirus species, and also indicates the possibility of engineering broad-spectrum antivirals by mimicking or intercepting L–Z interactions.

Discussion

The structures of L–Z complexes, coupled with previous and current biochemical evidence, reveal both conservation and variability

Fig. 5 | Catalysis of L polymerase and its modulation by Z protein. **a**, Structure of influenza virus polymerase at the pre-catalytic state, in closed conformation (PDB 6SZV)²¹. The catalytic motifs are shown as cartoons and the key residues are shown as sticks. Hydrogen bonds and salt bridges are represented by black dashed lines. The O3 atom of the -1 nucleotide and the α -phosphate are connected by a green dashed line to indicate the site for catalysis. For clarity, the sequence register numbers of RNA strands are labelled in parentheses, where the NTP substrate sits at the +1 position and the -1 nucleotide corresponds to the 3' end of primer strand. **b**, Structural model of LASV L in the open pre-catalytic conformation. The RNA strands and NTP substrate were modelled based on the influenza virus polymerase structure in **a**. The catalytic metal A (MeA) is offset from the catalytic position, which should be relocated to enable catalysis, as indicated by a black arrow. The supposed positions for the two metal ions ready for catalysis are represented by red dashed circles. **c**, Magnified view of catalytic motifs D and E in the LASV L structure. The Z protein is represented by a semi-transparent oval. **d**, Sequence context of motifs D (red) and E (yellow) in LASV L. The peptide fragments detected by mass spectrometry are underlined. **e**, Comparison of the deuterium uptake efficiencies of LASV L motifs in the presence or absence of Z proteins. The data are representative of three independent experiments using different protein preparations.

of L-Z interactions, providing an important basis for developing Z protein-based antiviral molecules. Contacting site 1 dominates the interaction and is highly conserved for all mammarenaviruses, in which the strictly conserved Trp residue plays a central role. The adjacent Phe residue is essential for MACV (and possibly all New World mammarenaviruses) but seems dispensable for Old World viruses, as evidenced by its replacement in LCMV and Lujo virus (LUJV) Z proteins (Extended Data Fig. 7c). This may explain the previous observation that LCMV Z was unable to inhibit the polymerase activity of MACV L in replicon cells¹⁴. We also confirmed this observation by in vitro polymerase assays and found that the LCMV Z protein showed no obvious inhibitory effect on MACV L polymerase but efficiently inhibited LASV L polymerase (Extended Data Fig. 10).

To analyse how a single conserved Trp35 (numbered by LASV Z) at site 1 stabilizes the L-Z interactions in LCMV and LUJV, we predicted the two complex structures by homology modelling, with the structure of LASV L-Z as the template (Extended Data Fig. 10). In both structures, a substitution of p.Met694Leu (numbered by LASV L) occurs relative to LASV L. The non-polar side chain of Leu694 forms additional contacts with the aromatic ring of Trp35 in the Z protein. Also, the p.Val650Phe substitution in LUJV L may contribute additional π - π interactions with Trp35 (Extended Data Fig. 10). Thus, LCMV and LUJV L polymerases may harbour more favourable hydrophobic molecular contexts to accommodate the strictly conserved Trp of the Z protein, enabling efficient binding in the absence of Phe36.

The residue Arg36 of MACV Z is also important for its inhibitory activity, which is conserved for all New World

Fig. 6 | Cross-reactivity of LASV and MACV L and Z proteins. **a,b**, Binding kinetics of MACV (**a**) and LASV (**b**) Z proteins to the opposite heterologous L polymerases. The data are representative of three independent experiments using different protein preparations. **c,d**, Cross-inhibition of MACV (**c**) and LASV (**d**) Z proteins to the opposite heterologous L proteins. Each data point represents the mean value of three independent experiments using different protein preparations. The error bars indicate the s.d. of each point.

arenaviruses (Extended Data Fig. 7). Similar effects have also been observed for TCRV—a member of the New World viral group²². In addition, three conserved hydrophobic residues, Tyr29, Leu41 and Tyr48 (numbered by LASV Z), were suggested as crucial for TCRV L–Z interactions²². These residues, which are not surface exposed or accessible for L polymerase, form a hydrophobic patch to stabilize the loop conformation in which the critical Arg residues (Extended Data Fig. 7e). Therefore, all of these data demonstrate a universal function of the Z protein for regulating polymerase activities, and also reveal both conserved and group-specific molecular determinants underpinning these interactions resulting from divergent evolution of Old World and New World mammarenaviruses.

Importantly, both the N- and C-terminal arms of the Z protein remain highly flexible when bound to the L polymerase, which offers an opportunity for the Z protein to simultaneously interact with multiple binding partners. One of the key roles of the Z protein is to govern the packaging of genomic segments into nascent viral particles. This may require the Z protein to be able to interact with the plasma membrane, via its N-terminal arm²⁵, when bound to the vRNP complex. The structures of L–Z complexes suggest that the Z protein may serve as an adaptor to recognize the functional vRNP by interacting with the resident L polymerase via its central LBD, and the flexible N-terminal arm may reach out for membrane targeting. This mechanism would ensure the precise packaging of viral genomic segments, as the interaction between L polymerase and vRNA is sequence specific²⁶, thus avoiding the production of defective virions.

Viral transcription and replication modulation by matrix proteins might be a common theme for negative-sense RNA viruses, which have also been observed in influenza virus, rabies virus, vesicular stomatitis virus and measles virus^{27–30}. The self-regulation of viral replication is important for balancing the propagation of progeny virions and host immunity. Interrupting the feedback loop of this regulation might be a feasible strategy to invoke and boost host immunity for viral clearance. Considering the conservation of contacting site 1 for arenavirus L–Z interaction, this motif of the Z

protein provides a promising molecular entity for further engineering to develop broad-spectrum antiviral drugs.

Methods

Cell lines. Insect SF9 cells (11496015; Invitrogen) were used to prepare the recombinant baculoviruses, and polymerase proteins were expressed in High Five cells (B85502; Invitrogen). Both cell lines were purchased from Invitrogen, routinely maintained in our laboratory and tested negative for *Mycoplasma* contamination. They were not experimentally authenticated in this study and are not included in the commonly misidentified cell line list of the International Cell Line Authentication Committee.

Protein production. Both LASV (strain G3278-SLE-2013; AIT17397.1) and MACV (strain Carvalho; AIG51560.1) L proteins were expressed using the Bac-to-Bac expression system, and purified by tandem affinity chromatography and size-exclusion chromatography (SEC) as previously described¹⁶. Briefly, the cells were harvested by centrifugation at 48 h post-infection and lysed by sonication. Cell lysates were clarified by centrifugation and ammonium sulfate was added at a final concentration of 0.5 g ml⁻¹ to precipitate the target protein. The precipitant was collected, resuspended and then applied to two-step affinity purification using HisTrap HP and StrepTrap HP (GE Healthcare) columns. The eluted sample was further purified by SEC using a Superose 6 10/300 GL (GE Healthcare) column equilibrated with a buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 5% (vol/vol) glycerol and 1 mM Tris(2-carboxyethyl) phosphine (TCEP).

The coding sequences of the LASV (AIT17396.1) and MACV (NP_899214.1) Z proteins were synthesized and codon-optimized for the *Escherichia coli* expression system. The sequence was fused with an N-terminal maltose-binding protein tag and a C-terminal 2xStrep tag and incorporated into the pET-21a expression vector. Recombinant proteins were expressed with the *E. coli* BL21 (DE3) strain at 16°C overnight in the presence of 0.5 mM isopropyl β-D-1-thiogalactopyranoside. The medium was supplemented with 0.1 mM ZnSO₄ to enhance the yield of Z proteins. Cells were harvested by centrifugation and homogenized with a high-pressure cell disrupter (JNBIO). The supernatant was collected by centrifugation and passed through a 0.22-μm cut-off filter to remove the insoluble debris. Target protein was captured by Ni-NTA affinity chromatography using a HisTrap HP (GE Healthcare) column, and further purified by SEC using a Superdex 200 10/300 GL (GE Healthcare) column equilibrated with a buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM TCEP. To remove the maltose-binding protein tag, tobacco etch virus protease was supplemented with a molar ratio of enzyme:substrate = 1:50 and incubated for 2 h at room temperature. The cleaved product was further purified by anion-exchange chromatography using a RESOURCE Q (GE Healthcare) column, in which the target protein flowed through the column. An additional round of SEC was performed for the

non-tagged protein using a Superdex 75 10/300 GL (GE Healthcare) column. The final product was analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) and concentrated with an Amicon Ultra Centrifugal concentrator (Millipore).

Cryo-EM sample preparation and data collection. To prepare L–Z protein complexes, LASV/MACV L polymerase was incubated with the corresponding Z protein with a molar ratio of L:Z = 1:3 at 4 °C for 1 h. The mixture was resolved by SEC using a Superose 6 10/300 GL (GE Healthcare) column equilibrated with 20 mM Tris–HCl (pH 7.5), 500 mM NaCl, 5% (vol/vol) glycerol and 1 mM TCEP. The complex fractions were concentrated to 5.0 mg ml^{−1} for cryo-EM sample preparation. Before vitrification, protein samples were diluted with a glycerol-free buffer to reduce the concentration of glycerol to 1% (vol/vol). An aliquot of 3 µl sample (1.0 mg ml^{−1}) was applied to a glow-discharged Quantifoil 1.2/1.3 holey carbon grid, which was blotted for 2.5 s with a humidity of 100% and then plunge frozen with an FEI Vitrobot Mark IV. Cryogenic specimens were loaded onto an FEI Titan Krios transmission electron microscope for data collection. The microscope was operated at 300 kV and a post-column GIF Quantum energy filter (Gatan) was used with a slit width of 20 eV. Cryo-EM data were automatically collected using SerialEM software (<http://bio3d.colorado.edu/SerialEM/>). Images were recorded with a Gatan K2 Summit detector using the super-resolution counting mode with a calibrated pixel size of 1.36 or 1.08 Å for LASV or MACV L–Z complexes, respectively. The exposure was performed with a dose rate of 10 e[−] pixel^{−1} s^{−1} and an accumulative dose of ~60 e[−] Å^{−2} for each image, which was fractionated into 30 movie frames. The final defocus ranges of the datasets were approximately −1.3 to −3.5 and −1.4 to −3.2 µm for the LASV and MACV L–Z complexes, respectively.

For the MACV L–Z–vRNA and LASV L–Z_{p.Phe36Ala}–vRNA complexes, the pre-assembled L–Z complexes were buffer-exchanged into 20 mM Tris–HCl (pH 7.5), 200 mM NaCl, 2% (vol/vol) glycerol and 1 mM TCEP and mixed with the 3' vRNA at a molar ratio of protein:RNA = 1:1.5. After 4 h incubation on ice, the sample (0.5 mg ml^{−1}) was directly subjected to vitrification. The cryogenic specimens were prepared using the Ni–Ti foil grids (1.2/1.3 spacing) following the same protocol as for the L–Z complexes. The two datasets were collected on a 200 kV Talos Arctica (Thermo Fisher Scientific) transmission electron microscope equipped with a post-column GIF Quantum energy filter (Gatan), which was used with a slit width of 20 eV. Cryo-EM data were automatically collected following the beam-image shift imaging scheme³¹. The images were recorded with a Gatan K2 Summit detector using the super-resolution counting mode with a calibrated pixel size of 1.00 Å at the specimen level. Each image was exposed with a dose rate of 10 e[−] pixel^{−1} s^{−1} and an accumulative dose of ~60 e[−] Å^{−2}, which was fractionated into 32 movie frames. The final defocus ranges were −0.9 to −2.4 and −1.1 to −2.6 µm for the MACV and LASV L–Z–vRNA complexes, respectively.

Image processing. The movie frames were aligned using MotionCor2 (ref. ³²) and the initial contrast transfer function (CTF) values for each micrograph were estimated with CTFFIND4.1 (ref. ³³). The images with an estimated resolution limit worse than 5 Å were discarded. An initial subset of ~10,000 particles were automatically selected from ten micrographs using Gautomatch (by K. Zhang), which was subjected to two-dimensional (2D) classification with RELION-3.0 (ref. ³⁴) to generate templates for auto-picking against the entire dataset. All subsequent classification and reconstruction steps were performed with RELION-3.0.

For the LASV L–Z dataset, approximately 1,200,000 particles were picked from ~1,500 micrographs. After two rounds of extensive 2D classification, a subset of ~859,000 particles were selected and subjected to 3D classification, with the density map of apo LASV L (Electron Microscopy Data Bank; EMD-0706) low-pass filtered to 60 Å resolution as the reference. Among the six 3D classes, a dominant class containing ~46% of total particles was identified, which displayed clear features of secondary structural elements. These particles were subjected to 3D refinement, which yielded a reconstruction at 4.4 Å resolution. To further improve the resolution, dose-weighted images were generated by MotionCor2 (ref. ³²) with the first two frames and last 13 frames discarded for each stack, resulting in a reduced dose of 30 e[−] Å^{−2}. In addition, CTF refinement was performed to correct the local CTF values of each particle. An additional round of 3D refinement was performed to obtain the final density map at 3.9 Å resolution estimated by the gold-standard Fourier shell correlation cut-off value of 0.143 (Extended Data Fig. 1). The processing of the dataset for LASV L bound to vRNA and the p.Phe36Ala Z protein mutant was quite straightforward, following the same procedure. A total of 2,732 micrographs were collected for this dataset, from which ~2,032,000 initial particles were picked. After multiple rounds of 2D classification and two consecutive rounds of 3D classification, a clean subset of 127,868 good particles were isolated, which led to a reconstruction at 3.4 Å resolution (Extended Data Fig. 8).

The MACV L–Z complex dataset was processed similarly. A total of ~987,000 initial particles were automatically picked from ~1,600 micrographs. After three rounds of 2D classification, approximately 535,000 good particles remained, which were subjected to 3D classification using the density map of apo MACV L (EMD-0707) low-pass filtered to 60 Å resolution as the initial model. A single dominant class was identified, accounting for 69% of the input particles, which

was used to calculate the final density map at 3.4 Å resolution by applying dose weighting and CTF refinement as described above (Extended Data Fig. 2). During 2D classification, we noticed that a small fraction of particles (~36,000) displayed features of L homodimerization, similar to the previous observations for apo or promoter-bound MACV L polymerase¹⁶. These particles led to a reconstruction at 4.1 Å by applying C2 symmetry. However, the orientation bias limited the quality of the density map. This map was not used for high-resolution information analysis. In this structure, two copies of Z protein bind to an L homodimer, with an identical contacting interface, as observed in the monomeric complex (Extended Data Fig. 2i).

For the MACV L–Z–vRNA complex dataset, a total of ~4 million particles were automatically picked from 5,754 micrographs. After two rounds of 2D classification, approximately 1.1 million monomeric particles and 0.5 million dimeric particles were isolated, which were individually processed in the subsequent steps. The monomeric particles were subjected to two rounds of 3D classification, and a clean subset of ~180,000 particles were selected. These particles were used to calculate a density map at 3.1 Å resolution (Extended Data Figs. 4 and 5). For the dimeric particles, approximately 164,000 good particles were selected after 3D classification by applying C2 symmetry, which generated a reconstruction at 3.4 Å resolution after 3D refinement. In this structure, the density for the CTD of the L protein was degraded due to symmetry miss-match and conformational heterogeneity. To improve the resolution, symmetry expansion was applied to facilitate reconstruction within an asymmetric unit. The expanded particles were recentred to the centre of mass of one protomer, and signal subtraction was performed to remove the density of the other protomer. The mask also covered partial density of the other protomer to keep the dimeric interface intact. The resulting monomeric particles were subjected to 3D refinement without applying symmetry, which improved the resolution to 3.2 Å. In this density map, the CTD of the L protein still could not be well defined. To better resolve this region, further signal subtraction and recentring were performed for the CTD. These subtracted particles were 3D-classified without alignment and restrained by a soft mask covering the density of the target region. Among the eight 3D classes, a distinguished class with ~110,000 particles was identified, which resulted in a reconstruction at 3.8 Å resolution (Extended Data Fig. 4c). The body of the monomeric map within the dimeric complex after symmetry expansion (3.2 Å) and the locally reconstructed density of the CTD (3.8 Å) were combined to generate a composite map of the dimeric complex for model building (Fig. 1d). We also performed reconstruction for these dimeric particles without applying symmetry, to analyse the potential differences in conformations or RNA binding of the two protomers. This strategy generated a density map at 3.6 Å resolution, which revealed that both copies of L proteins were bound by the 3' vRNA and in highly similar conformations (Supplementary Fig. 1). For the monomeric map at 3.1 Å resolution, an atomic model was built in to facilitate comparative analysis. Both the dimeric and monomeric forms displayed highly similar features for interactions with vRNA and Z protein (Supplementary Fig. 1).

Local resolution was assessed with ResMap³⁵. The core region of the LASV L–Z complex reached 3.0 Å resolution and the LASV L–Z_{p.Phe36Ala}–vRNA, MACV L–Z and L–Z–vRNA complexes were better resolved, with most regions extending to sub-3.0 Å local resolution. The cap-binding domain of L polymerase in the MACV L–Z–vRNA complex was resolved at ~4.0 Å local resolution, which clearly defined the main chain of β-strands (Extended Data Figs. 4h and 5c).

Model building and refinement. The structure of the apo LASV or MACV L protein was docked into the corresponding maps using CHIMERA³⁶. Additional density was identified for the Z protein in each structure, which was well fitted with the crystal structure of LASV Z protein (PDB 5172). The model was manually corrected for local fit in COOT³⁷ and the sequence register was updated based on alignment. The density for the CTD was missing in the L–Z complexes and LASV L–Z_{p.Phe36Ala}–vRNA complex, and the entire pendant insertion domain (residues 932–1085) of LASV L was not resolved, whereas roughly two-thirds of this region could be modelled with confidence for MACV L. The CTD was well defined in the MACV L–Z–vRNA complex in dimeric configuration. This region was modelled by manually main-chain tracing assisted by secondary structure prediction using the PsiPred server³⁸. The connecting loop between the PB2-N-like region and CTD was well defined and most of the side chains in the mid-domain could be identified. The main chain trace of the cap-binding domain could be well recognized but most of the side chains could not be confidently assigned. For the Z proteins, only the central domain was observed in all complexes. The models were refined against corresponding maps in real space using PHENIX³⁹, in which the secondary structural restraints and Ramachandran restraints were applied. The stereochemical quality of each model was assessed using MolProbity⁴⁰.

The representative density and atomic models are shown in Extended Data Figs. 1h, 2h, 5c and 8g. The statistics for image processing and model refinement are summarized in Supplementary Tables 1 and 2. The structural figures were rendered using CHIMERA³⁶, CHIMERAX⁴¹ or PyMOL (<https://pymol.org/>).

BLI binding assay. The binding affinities of the MACV and LASV Z proteins to L polymerases were measured using the Octet RED96 system (FortéBio). All of the experiments were performed at 25 °C in a buffer containing 50 mM Tris–HCl

(pH 7.5), 500 mM NaCl, 5% (vol/vol) glycerol, 1 mM TCEP, 0.1% bovine serum albumin and 0.05% Tween 20. Streptavidin biosensors were pre-equilibrated in buffer for at least 10 min before applying samples. Biotinylated Z proteins or mutants (100 nM) were loaded onto streptavidin biosensors for 120 s to immobilize. The sensor was then blocked with biocytin before applying analytes. Serially diluted L polymerase solutions were then applied to analyse the binding kinetics. The interference signals from the immobilized protein with buffer and the uncoated biosensors with analytes at corresponding concentrations were recorded as two sets of controls, which were subtracted from the raw binding curves for correction. The normalized data were used to calculate the binding constant values for quantifying the binding affinities using Octet Data Analysis Software.

In vitro polymerase activity assays. Since the 5' vRNA stimulates the replication activity but inhibits transcription of arenavirus polymerase¹⁶, we used de novo replication (in the presence of 5' and 3' vRNA) and cap-dependent transcription (with 3' vRNA only) reactions to test the inhibitory activities of Z protein and mutants. For the replication assay, the 3' vRNA template strand (5'-GCCUAGG AUCCACUGUGCG-3') and 5' vRNA (5'-CGCACCGGGGAUCCUAGG-3') were separately denatured at 65 °C for 5 min and then immediately cooled on ice for 5 min. The L protein (0.4 µM) was mixed with the 3' vRNA (1.0 µM) and 5' vRNA (1.0 µM) in the reaction buffer containing 50 mM Tris-HCl (pH 7.0), 40 mM NaCl, 10 mM KCl, 5 mM MnCl₂, 1 mM dithiothreitol and 0.1 mg ml⁻¹ bovine serum albumin. After 15 min incubation at 25 °C, Z proteins (wild type or mutants) were added at the specified concentrations and incubated for an additional 15 min. To initiate the reaction, 0.2 µCi µl⁻¹ [α -³²P]GTP, 1.0 mM ATP/UTP/CTP, 100 µM GTP and 0.5 U µl⁻¹ RNasin were added to the mixture, which was incubated at 30 °C for 3 h. For the transcription assay, the experiment was conducted with a similar procedure except that 3 µM capped RNA (5'-m⁷GpppAAUCCG-3'; TriLink) was supplemented to replace the 5' vRNA.

The reactions were stopped by adding an equal volume of formamide. The products were denatured by heating to 98 °C for 20 min, then resolved by 20% polyacrylamide gels containing 9 M urea run with 0.5× TBE buffer. The gels were visualized by exposure on a storage phosphor screen and read with a Typhoon scanner. To quantify the RNA products, the intensity of each band was integrated using ImageJ software. Statistical significance was determined by one-way analysis of variance or Student's *t*-test, where appropriate, using the SPSS program. The histograms and plots were prepared with Origin or GraphPad Prism software.

HDX-MS assay. The stock solutions of LASV L and Z proteins were prepared in a buffer containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl and 2% (vol/vol) glycerol, dissolved with H₂O. To prepare the L-Z complex, the L polymerase was mixed with Z protein with a molar ratio of L:Z = 1:2.5 and incubated on ice for 2 h. The hydrogen-deuterium exchange reaction was initiated by diluting the stock solution of LASV L (5.0 mg ml⁻¹), with or without Z protein, tenfold with a D₂O dissolved buffer consisting of 20 mM Tris-DCl (pD = 7.1) and 500 mM NaCl. The exchange reaction was conducted at room temperature. Aliquot samples were taken at four time points during the exchange reaction (60 s, 90 s, 300 s and 24 h) for quantification by mass spectrometry. Each sample was mixed with an equal volume of quench buffer (4 M guanidinium chloride, 200 mM citric acid, 100 mM TCEP (pH 2.0) dissolved in H₂O) to stop the exchange reaction. The protein was then digested with pepsin for 10 min on ice and the resulting peptides were injected into a tandem high-performance liquid chromatography (Thermo Fisher Scientific; Ultimate 3000) and mass spectrometry (Thermo Fisher Scientific; Q Exactive Orbitrap) system for analysis. The data were collected and processed with Xcalibur and HDEaminer software, respectively. No correction was made for back exchange and all of the results are reported as relative deuterium levels. The samples taken at the 24 h time point during the exchange reaction were thought to be saturated by deuterium, which was used to normalize the relative exchange efficiencies at other time points.

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

Data availability

The cryo-EM density maps and atomic coordinates have been deposited in the EMD and PDB, respectively. The accession numbers are as follows: LASV L-Z complex (PDB 7CKL and EMD-30387); MACV L-Z complex in monomeric form (PDB 7CKM and EMD-30388) and dimeric form (PDB 7ELB and EMD-31179); LASV L bound to 3' vRNA and the p.Phe36Ala Z protein mutant (PDB 7ELA and EMD-31178); and MACV L-Z-vRNA complex in monomeric form (PDB 7ELC and EMD-31180) and dimeric form (PDB 7EL9 and EMD-31177). Source data are provided with this paper. All other data are available from the authors on reasonable request.

Received: 8 December 2020; Accepted: 5 May 2021;
Published online: 14 June 2021

References

- Charrrel, R. N. & de Lamballerie, X. Arenaviruses other than Lassa virus. *Antivir. Res* **57**, 89–100 (2003).
- Gao, G. F. From “A” IV to “Z”IKV: attacks from emerging and re-emerging pathogens. *Cell* **172**, 1157–1159 (2018).
- Patterson, M., Grant, A. & Paessler, S. Epidemiology and pathogenesis of Bolivian hemorrhagic fever. *Curr. Opin. Virol.* **5**, 82–90 (2014).
- McCormick, J. B., Webb, P. A., Krebs, J. W., Johnson, K. M. & Smith, E. S. A prospective study of the epidemiology and ecology of Lassa fever. *J. Infect. Dis.* **155**, 437–444 (1987).
- Akpede, G. O., Asogun, D. A., Okogbenin, S. A. & Okokhere, P. O. Lassa fever outbreaks in Nigeria. *Expert Rev. Anti Infect. Ther.* **16**, 663–666 (2018).
- McKee, K. T. Jr., Oro, J. G., Kuehne, A. I., Spisso, J. A. & Mahlandt, B. G. Candid No. 1 Argentine hemorrhagic fever vaccine protects against lethal Junin virus challenge in rhesus macaques. *Intervirology* **34**, 154–163 (1992).
- McCormick, J. B. et al. Lassa fever. Effective therapy with ribavirin. *N. Engl. J. Med.* **314**, 20–26 (1986).
- Buchmeier, M. J., de la Torre, J. C., Peters, C. J. & Bowen, M. D. Arenaviridae: the viruses and their replication. *Fields Virol.* **2**, 1791–1827 (2007).
- Lopez, N., Jacamo, R. & Franze-Fernandez, M. T. Transcription and RNA replication of Tacaribe virus genome and antigenome analogs require N and L proteins: Z protein is an inhibitor of these processes. *J. Virol.* **75**, 12241–12251 (2001).
- Meyer, B. J., de la Torre, J. C. & Southern, P. J. Arenaviruses: genomic RNAs, transcription, and replication. *Curr. Top. Microbiol. Immunol.* **262**, 139–157 (2002).
- Fan, L., Briese, T. & Lipkin, W. I. Z proteins of New World arenaviruses bind RIG-I and interfere with type I interferon induction. *J. Virol.* **84**, 1785–1791 (2010).
- Cornu, T. I. & de la Torre, J. C. RING finger Z protein of lymphocytic choriomeningitis virus (LCMV) inhibits transcription and RNA replication of an LCMV S-segment minigenome. *J. Virol.* **75**, 9415–9426 (2001).
- Perez, M., Craven, R. C. & de la Torre, J. C. The small RING finger protein Z drives arenavirus budding: implications for antiviral strategies. *Proc. Natl Acad. Sci. USA* **100**, 12978–12983 (2003).
- Kranzusch, P. J. & Whelan, S. P. Arenavirus Z protein controls viral RNA synthesis by locking a polymerase–promoter complex. *Proc. Natl Acad. Sci. USA* **108**, 19743–19748 (2011).
- Jacamo, R., Lopez, N., Wilda, M. & Franze-Fernandez, M. T. Tacaribe virus Z protein interacts with the L polymerase protein to inhibit viral RNA synthesis. *J. Virol.* **77**, 10383–10393 (2003).
- Peng, R. et al. Structural insight into arenavirus replication machinery. *Nature* **579**, 615–619 (2020).
- Lehmann, M. et al. Role of the C terminus of Lassa virus L protein in viral mRNA synthesis. *J. Virol.* **88**, 8713–8717 (2014).
- Morin, B. et al. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. *PLoS Pathog.* **6**, e1001038 (2010).
- Hastie, K. M. et al. Crystal structure of the oligomeric form of Lassa virus matrix protein Z. *J. Virol.* **90**, 4556–4562 (2016).
- Cornu, T. I. & de la Torre, J. C. Characterization of the arenavirus RING finger Z protein regions required for Z-mediated inhibition of viral RNA synthesis. *J. Virol.* **76**, 6678–6688 (2002).
- Wandzik, J. M. et al. A structure-based model for the complete transcription cycle of influenza polymerase. *Cell* **181**, 877–893 (2020).
- Loureiro, M. E. et al. Molecular determinants of arenavirus Z protein homo-oligomerization and L polymerase binding. *J. Virol.* **85**, 12304–12314 (2011).
- Wilda, M., Lopez, N., Casabona, J. C. & Franze-Fernandez, M. T. Mapping of the Tacaribe arenavirus Z-protein binding sites on the L protein identified both amino acids within the putative polymerase domain and a region at the N terminus of L that are critically involved in binding. *J. Virol.* **82**, 11454–11460 (2008).
- Peng, Q. et al. Structural basis of SARS-CoV-2 polymerase inhibition by favipiravir. *Innovation* **2**, 100080 (2020).
- Perez, M., Greenwald, D. L. & de la Torre, J. C. Myristoylation of the RING finger Z protein is essential for arenavirus budding. *J. Virol.* **78**, 11443–11448 (2004).
- Kranzusch, P. J. et al. Assembly of a functional Machupo virus polymerase complex. *Proc. Natl Acad. Sci. USA* **107**, 20069–20074 (2010).
- Watanabe, K., Handa, H., Mizumoto, K. & Nagata, K. Mechanism for inhibition of influenza virus RNA polymerase activity by matrix protein. *J. Virol.* **70**, 241–247 (1996).
- Iwasaki, M. et al. The matrix protein of measles virus regulates viral RNA synthesis and assembly by interacting with the nucleocapsid protein. *J. Virol.* **83**, 10374–10383 (2009).
- Finke, S., Mueller-Waldeck, R. & Conzelmann, K. K. Rabies virus matrix protein regulates the balance of virus transcription and replication. *J. Gen. Virol.* **84**, 1613–1621 (2003).

30. Clinton, G. M., Little, S. P., Hagen, F. S. & Huang, A. S. The matrix (M) protein of vesicular stomatitis virus regulates transcription. *Cell* **15**, 1455–1462 (1978).
31. Wu, C., Huang, X., Cheng, J., Zhu, D. & Zhang, X. High-quality, high-throughput cryo-electron microscopy data collection via beam tilt and astigmatism-free beam-image shift. *J. Struct. Biol.* **208**, 107396 (2019).
32. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
33. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
34. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
35. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
36. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
37. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
38. Buchan, D. W. A. & Jones, D. T. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* **47**, W402–W407 (2019).
39. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
40. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).
41. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
42. Volpon, L., Osborne, M. J., Capul, A. A., de la Torre, J. C. & Borden, K. L. Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E. *Proc. Natl Acad. Sci. USA* **107**, 5441–5446 (2010).

Acknowledgements

We thank all of the staff members at the Center for Biological Imaging (CBI), Institute of Biophysics (IBP), Chinese Academy of Sciences (CAS) and Cryo-EM Center of the Southern University of Science and Technology (SUST) for assistance with data collection. We are grateful to G. Wang (Core Facilities at the School of Life Sciences, Peking University) and staff at the Electron Microscopy department of the State Key Laboratory of Membrane Biology, Institute of Zoology (IOZ), CAS, Beijing for technical

support with electron microscope operation. We thank the Center of Protein Analysis Technology, Tsinghua University for assistance with HDX-MS analysis. This study was supported by the Strategic Priority Research Program of CAS (XDB29010000 to Y.S.), National Science and Technology Major Project (2018ZX10101004 to Y.S.), External Cooperation Program of CAS (153211KYSB20160001 to Y.S. and G.F.G.), Advanced Customer Cultivation Project of the Wuhan National Biosafety Laboratory, CAS (2018ACCP-MS11 to Y.S.) and National Natural Science Foundation of China (NSFC) (81871658 to Y.S.). M.W. is supported by the NSFC (81802007). R.P. is supported by the Young Elite Scientist Sponsorship Program (YESS) of the China Association for Science and Technology (CAST; 2018QNRC001). Y.S. is also supported by the Excellent Young Scientist Program from NSFC (81622031), and the Youth Innovation Promotion Association of CAS (Y201921).

Author contributions

Y.S., X.X. and R.P. conceived of the study idea. X.X., Q.P., M.W. and Y.X. purified the protein samples. X.X., Q.P., M.W., Y.T. and X.H. conducted the polymerase activity assays. R.P., Q.P., S.L. and P.W. prepared the cryo-EM specimens and collected the data. R.P. performed the image processing and reconstruction. R.P. and J.Q. built the atomic models. X.X., R.P., Q.P., M.W. and Y.S. analysed the structures. X.T. and H.D. performed the HDX-MS assay. R.P., X.X., Q.P. and Y.S. wrote the initial draft. J.Z., P.W. and G.F.G. contributed to the final version of the manuscript. All authors participated in the discussion and editing and approved the manuscript. Y.S. supervised all of the research.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41564-021-00916-w>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41564-021-00916-w>.

Correspondence and requests for materials should be addressed to Y.S.

Peer review information *Nature Microbiology* thanks Robert Garry and the other, 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 www.nature.com/reprints.

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

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Extended Data Fig. 1 | Cryo-EM analysis of LASV L-Z complex. **a**, Size-exclusion chromatogram and SDS-PAGE profile of LASV L-Z complex. The data shown are representative results of two independent experiments using different protein preparations. **b**, A representative micrograph of LASV L-Z complex (one out of ~1,500 micrographs). The scale bar represents 50 nm. **c**, Gallery of 2D class average images of LASV L-Z particles. The scale bar represents 10 nm. Two rounds of 2D classification were performed during image processing. **d**, A brief workflow of cryo-EM image processing and reconstruction. **e**, Euler angle distribution of the final reconstruction. **f**, Local resolution distribution for the final density map. **g**, FSC curves for the reconstruction. **h**, Representative density and atomic models in selected regions.

Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Cryo-EM analysis of MACV L-Z complex. **a**, Size-exclusion chromatogram and SDS-PAGE profile of MACV L-Z complex. The data shown are representative results of two independent experiments using different protein preparations. **b**, A representative micrograph of MACV L-Z complex (one out of ~1,600 micrographs). The scale bar represents 50 nm. **c**, Gallery of 2D class average images of MACV L-Z particles. The scale bar represents 10 nm. Three rounds of 2D classification were performed during image processing. **d**, A brief workflow of cryo-EM image processing and reconstruction. **e**, Euler angle distribution of the final reconstruction. **f**, Local resolution distribution for the density map of monomeric L-Z complex. **g**, FSC curves for the monomeric reconstruction. **h**, Representative density and atomic models in selected regions. **i**, Density map of the dimeric form, (L-Z)₂, of MACV L-Z complex. The CTD could be visualized in this configuration, indicated by blue dashed ovals. The two copies of Z protein are highlighted with red dashed circles.

Extended Data Fig. 3 | Architecture of LASV/MACV L polymerases and their different conformational states. a, Schematic diagram for the domain architecture of LASV and MACV L polymerases. **b**, Structure of the full-length MACV L polymerase. The structure is shown in cartoons and colored by domains with the same scheme as in (a). The Z protein binding site is indicated by a semi-transparent oval. The CTD is flexibly linked to the body of L polymerase via a long linker loop (C-linker), highlighted by a black arrow. **c**, Comparison of the structures of LASV and MACV L-Z complexes in the presence or absence of 3'-vRNA. The pendant domain and CTD are not resolved in LASV L polymerase structures, indicated by dashed ovals. The α -ribbon and pendant domains are well resolved in MACV L-Z complex, which may block the entrance for 3'-vRNA template. In the presence of 3'-vRNA, the two domains become disordered due to potential conformational changes. The 5'-end portion of 3'-vRNA strand are not resolved in these structures, represented by black dashed lines. The putative template entrance is indicated by thick black arrows.

Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Cryo-EM analysis of MACV L-Z-vRNA complex. **a**, A representative micrograph of MACV L-Z complex (one out of >5,000 micrographs). The scale bar represents 50 nm. **b**, Gallery of 2D class average images of MACV L-Z-vRNA complex. The scale bar represents 10 nm. Two rounds of 2D classification were performed during image processing. Two types of complex were observed in the micrographs, the 'monomeric' (L-Z-vRNA) form and the 'dimeric' form (L-Z-vRNA)₂. **c**, A brief workflow of cryo-EM image processing and reconstruction. **d**, Euler angle distribution profile of the monomeric complex. **e**, Euler angle distribution profile of the asymmetric unit within the dimeric complex by reconstruction without applying symmetry. **f**, Euler angle distribution of the dimeric complex reconstructed with C1 symmetry. **g-i**, Local resolution distribution for the corresponding density maps.

Extended Data Fig. 5 | Quality assessment of the MACV L-Z-vRNA complex reconstruction. a-b, FSC curves for the monomeric and dimeric forms of MACV L-Z-vRNA complex. **c,** Representative density and atomic models in selected regions.

Extended Data Fig. 6 | Structural plasticity of Z protein. **a**, Predicted model of arenavirus polymerase in elongation state. This model is generated based on the elongation structure of influenza virus polymerase (PDB ID: 6QCT)²¹. The entry/exit tunnels for the template and product strands are indicated by arrows. The direction for the product strand to extent is indicated by red dashed arrows. The entire CTD of arenavirus polymerase is linked to the core RdRp via a highly flexible linker loop, denoted as the C-linker. The L polymerase is shown in cartoons and colored by domains. The Z protein is shown as surface model which binds at the palm domain of RdRp. **b**, Oligomeric structure of LASV Z protein (PDB ID: 5I72)⁹. The structure is visualized as ribbon models. The bound Zinc atoms are shown as spheres. **c**, Comparison of the L-binding footprint (red) and the contacting interface for Z self-oligomerization (pink). The two surfaces are partially overlapped. **d**, Conformational plasticity of Z protein in different structures/functional states. The inset shows the conformational variations of a key hairpin-loop (35-hairpin) important for interaction with L polymerase. The presented monomeric structure of Z is one of the conformers determined by NMR in solution (PDB ID: 2M1S)⁴².

Extended Data Fig. 7 | The conserved contacting site 1 for L-Z interactions. **a-b**, The two conserved W and F residues of Z insert into the hydrophobic cavity in LASV (**a**) and MACV (**b**) L polymerases. The protruding loop of Z proteins is shown in cartoons (colored in salmon) and the polymerases are shown in surface models (colored by domains). The key hydrophobic residues around the cavity are shown as sticks, for which the hydrophobic surfaces are highlighted in yellow. **c**, Sequence alignment of representative mammarenavirus Z proteins for the LBD in L-Z complexes, including the central RING domain (highlighted in yellow) and a small portion of the terminal arms. The sequences are numbered based on LASV Z. The two Zinc-finger motifs are connected by lines. The strictly conserved determinant residue W is highlighted by a blue arrowhead. Two other less conserved residues, F and R, are indicated by green arrowheads, which are crucial for binding L polymerases in NW arenaviruses, but not OW arenaviruses. **d**, Conservation analysis of mammarenavirus Z proteins. The two conserved hydrophobic residues critical for L binding in contacting site 1 and the two Zinc-finger motifs are shown as sticks. **e**, Three conserved residues, Y29, L41 and Y48 (numbered by LASV Z), were suggested critical for L-Z interactions by previous studies²². They are not directly involved in binding L polymerase, but are important for maintaining the conformation of the adjacent loop which contains a key interacting residue R conserved for all NW mammarenaviruses.

Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Cryo-EM analysis of LASV L bound to 3'-vRNA and Z F36A mutant. **a**, A representative micrograph of LASV L-Z F36A-vRNA complex (one out of ~2,700 micrographs). The scale bar represents 50 nm. **b**, Gallery of 2D class average images. The scale bar represents 10 nm. Three rounds of 2D classification were performed during image processing. **c**, A brief workflow of cryo-EM image processing and reconstruction. **d**, Euler angle distribution profile of the reconstruction. **e**, FSC curves for assessment of the reconstruction. **f**, Local resolution distribution of the density map. **g**, Representative density and atomic models in selected regions.

Extended Data Fig. 9 | Verification of the inhibitory efficacy of LASV Z F36A mutant. **a**, The sequences of RNA templates used for in vitro transcription and replication assays. The shorter template (Temp_1) corresponds to the conserved 3'-vRNA promoter of arenaviruses, which was used for most in vitro polymerase assays. The longer one (Temp_2) contains additional upstream sequence in the viral genome, which was used to better resolve the inhibitory effect of LASV Z F36A mutant. **b-c**, Inhibitory effects of LASV Z and mutant on the transcription and replication of L polymerase using a longer template RNA. The WT and mutant Z proteins were used at $1.5 \times IC_{50}$ concentrations, consistent with the experiments in Fig. 3, at which the WT protein displayed >90% inhibition for polymerase activities. The data shown are mean values (histograms) $\pm$ s.d. (error bars) of three independent experiments using different protein preparations. The significance of difference was tested by one-way analysis of variance and the P values were calculated for each group.

Extended Data Fig. 10 | Predicted models of L-Z interaction interface of LCMV and LUJV. a-b, The structures of LCMV (**a**) and LUJV (**b**) L bound with cognate Z proteins were predicted by homology modelling based on the structure of LASV L-Z complex. The LCMV and LUJV L-Z structures are colored by domains and overlaid to LASV L-Z complex (white). The key interacting residues at site 1 are shown as sticks and colored by atoms. For clarity, all residues are numbered by the sequences of LASV proteins. The predicted interaction between L649 and W35 is represented by red dashed lines. An additional π - π interaction between F650 and W35 may exist in the LUJV L-Z complex, indicated by a blue dashed line. **c-d,** Effects of LCMV Z protein on the activity of LASV (**c**) and MACV (**d**) L proteins. The LCMV Z protein was used at the same concentration as in Fig. 3d, corresponding to $1.5 \times IC_{50}$ concentration of LASV Z protein, at which the LASV Z protein displayed near complete inhibition for LASV L polymerase. The data shown are mean values (histograms) $\pm$ s.d. (error bars) of three independent experiments using different protein preparations. The significance of difference was tested by student's *t*-test (two sided) and the *P* values were calculated for each group.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒ ☐ A description of all covariates tested
- ☒ ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	SerialEM-3.7 was used for cryo-EM data collection, Octet RED96 system was used for BLI data collection and Xcalibur-4.3 was used for HDX-MS data collection.
Data analysis	MotionCor2-1.1.0; CTFFIND-4.1; Gautomatch-v0.56; RELION-3.0; ResMap-1.1.4; CHIMERA-1.11.1; CHIMERAX-1.0; COOT-0.8.1; PHENIX-1.11.1-2575; MolProbity-4.3.1; PyMOL-2.0; Octet Data Analysis Software-10.0; Image J-1.52a; OriginPro-2016; SPSS-24.0.0.0; HDExaminer-3.0; GraphPad prism 6

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The cryo-EM density maps and atomic coordinates have been deposited to the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB), respectively. The accession numbers are listed as follows: LASV L-Z complex (PDB-7CKL, EMD-30387), MACV L-Z complex in monomeric form (PDB-7CKM, EMD-30388) and dimeric form (PDB-7ELB, EMD-31179), LASV L bound to 3'-vRNA and Z F36A mutant (PDB-7ELA, EMD-31178), MACV L-Z-vRNA complex in monomeric form (PDB-7ELC, EMD-31180) and dimeric form (PDB-7EL9, EMD-31177). The uncropped gel/autoradiograph images and raw quantification data for main figures (Figs. 3, 5 and 6) and Extended Data figures (Extended Data Figs. 9 and 10) are provided as associated source data files. The source data for

Supplementary figures (Supplementary Figs. 2 and 3) are provided in the Supplementary Information files as stated in the legends. All other data are available from the authors on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. The cryo-EM data for each sample were collected for at least one day until the data generated interpretable maps with sufficient resolutions reported in this study.
Data exclusions	We performed 2D and 3D classification procedures for the initial particle images to discard the bad particles or classes that did not result in good 3D reconstructions. This procedure is standard in the EM field. The details for these procedures are described in the Methods section.
Replication	The size-exclusion chromatography assays, in vitro polymerase activity assays and hydrogen-deuterium exchange-mass spectrometry assays were performed in two or three independent replicates using different protein preparations, which all generated similar results. The related replicative results were used for quantification and statistical analyses, which are provided in the source data and described in the figure legends accordingly.
Randomization	This is not relevant to our study, because no grouping was needed.
Blinding	Investigators were not blinded to group allocation, as no grouping was needed for this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Insect cells (SF9 and HI5 cells, purchased from the Invitrogen Company) were used for protein expression using the baculovirus expression system.
Authentication	The cells were purchased and routinely maintained in our lab. They were not authenticated experimentally for these studies.
Mycoplasma contamination	All cell lines are tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	The insect cell lines we used are not listed in ICLAC. In addition, the cells themselves are not being studied, instead they are used as tools for recombinant protein expression.