PNAS
www.pnas.org

Supplementary Information for

The mosquito protein AEG12 displays both cytolytic and antiviral properties via a common lipid transfer mechanism

Alexander C.Y. Foo[1], Peter M. Thompson[1], Shih-Heng Chen[1], Ramesh Jadi[2], Brianna Lupo[1], Eugene F. DeRose[1], Simrat Arora[1], Victoria C. Placentra[1], Lakshmanane Premkumar[2], Lalith Perera[1], Lars C. Pedersen[1], Negin Martin[1], Geoffrey A. Mueller[1*]

1- Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences
2- Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA.

*Geoffrey A. Mueller

Email: Geoffrey.Mueller@nih.gov

This PDF file includes:

Supplementary text
Figures S1 to S6 (not allowed for Brief Reports)
Tables S1 to S1 (not allowed for Brief Reports)
Legends for Movies N/a
Legends for Datasets N/a
SI References

Other supplementary materials for this manuscript include the following:

Movies N/a
Datasets N/a

Supplementary Information Text

Supplemental Materials and Methods

AEG12 Constructs and X-ray Crystallography: AEG12 in this paper refers specifically to AAEL013577-PB residues 20-207, representing a single insect MA domain repeat without the N-terminal signal peptide (1). AEG12 was expressed in BL21 E. coli as a fusion with an N-terminal Glutathione-S-Transferase (GST) tag separated by a TEV cut site and a three-alanine linker. Un-labelled and uniformly ^{15}N-labeled AEG12 was cultured, and expressed as previously described for Bla g 1 (2). For NMR studies, uniformly ^{15}N and per-deuterated ^{13}C-^{15}N-lebeled AEG12 for NMR were expressed using modified minimal M9 media as previously described.(2, 3) Protein concentrations were assessed using a bicinchoninic acid assay kit (Pierce). A theoretical binding stoichiometry of 8 fatty acids or 4 diacyl chain ligands was assumed for all studies. For X-ray crystallography studies, residues 24-207 were instead transferred into a pMALX(E) vector (New England Biolabs) and expressed as an MBP fusion in Rosetta(DE3). Cells were grown in TB media and induced with 0.4 mM IPTG. AEG12 was purified from the resulting cells using an amylose resin column (New England Biotechnology) and a Superdex200 26/60 column (GE) for crystallization.

Crystals of MBP(E)-AEG12 were obtained at 20 °C by sitting drop vapor diffusion by mixing 325 nl concentrated protein with 325 nl of reservoir solution containing 50 mM MES pH 6.5, 100 mM NaCl, 200 mM $MgCl_2$ and 20% PEG1000. For data collection, 1µl of cryo solution consisting of 50 mM MES pH 6.5, 125 mM NaCl, 200 mM $MgCl_2$, 22% PEG1000, 5 mM Maltose, and 15% ethylene glycol was added to the crystallization drop. The crystal was then transferred to 100% cryo-solution for 30 seconds then flash frozen in liquid nitrogen. The 2.1 Å data set was collected on a Rigaku MicroMax007HF generator equipped with a Saturn944 CCD detector. The 1.95 Å data set was collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory at a wavelength of 1.0 Å. All data were processed using HKL3000 and HKL2000, respectively (4, 5), Molecular replacement of the MBP(E)-AEG12 structure was carried out with the 2.1 Å data in Phenix (6) using Phaser (7) and coordinates PDB ID code = 5T05 for the MBP and sculpted coordinates PDB ID code 4JRB of Bla g 1(8, 9). After minimal refinement, the individual MBP and AedesZ components of the model were used to solve the molecular replacement for the 1.95 Å data set while maintaining the same Rfree test set. The crystal structure was refined using multiple cycles of model building in Coot with positional and B-factor refinement in Phenix (6, 10). Structure was determined to have good geometry statistics by Molprobity (11). Data statistics are reported in Table S1.

Computational Studies: The X-ray structure of AEG12 was used to generate the initial structure for MD simulations. 1-4 DSPC ligands or 7-12 oleate ligands were manually introduced and energy-minimized using the Amber 16.(12) Each cargo-bound or Apo-AEG12 model was solvated in a box containing over 23000 water molecules and counterions (Na^+ and Cl^-) were added. Prior to equilibration, all systems were subjected to 1) 100-ps belly dynamics runs with fixed peptide, 2) minimization, 3) fixed protein, low temperature constant pressure dynamics 4) minimization, 5) step-wise slow heating molecular dynamics at constant volume, and 6) constant volume unconstrained molecular dynamics for 10 ns. Following an extra 20 ns equilibration step, unconstrained MD trajectories were continued at 300 K under constant pressure for 120 ns (DSPC and oleate 1-6) or 500 ns (oleate) with a 1 fs timestep, using the PMEMD module of Amber.16 to accommodate long range interactions. Partial charges on lipids were calculated using Gaussian.09.(13) at the B3LYP/6-31g level. The force field paramaters were taken from the FF14SB force field for the protein and the lipid and phospholipid parameters were from lipid14 in the AMBER.16 package.

Cell Lysis: Human erythrocytes from either Hispanic or African-American subjects were obtained from ZenBio, packaged with either EDTA or ACD anti-clotting agents. Samples were resuspended in RBC buffer (100 mM NaCl, 10 mM KCl, 25 mM K_2PO_4, 15 mM glucose, 10 mM ascorbate, and 25U heparin), rinsed 3x in PBS and resuspended to a final haematocrit of 25%.

HEK-293 (freestyle) cells were grown using the Freestyle expression medium (Thermo-fisher). SF9 (ATCC CRl-1711) cells were cultured in Sf900III, 1X Anti-Anti, 5% Fetal bovine serum, and Vero (ATCC CCL-81) cells were cultured in Dulbecco's modified Eagle medium,10%FBS, 4 mM L-Glutamine, and 1mM Sodium Pyruvate. Cells were pelleted at 600x g, rinsed 3x in PBS, and resuspended to produce a final concentration

of $1x10^6$ mL^{-1} for all assays. Cultures of the diploid JFS1427 yeast strain were inoculated from single colonies and incubated at 23°C overnight in YPDA broth. Cells were washed 2x in PBS, diluted to a final concentration of $1x10^6$ mL^{-1} for all assays. *E. coli* BL21 DE3 cells containing a kanamycin resistance pET 9a vector were grown to an OD_{600} of 1.0, pelleted, rinsed 3x with PBS and resuspended to yield a final concentration of $1x10^7$ mL^{-1} for all assays.

Mass Spectrometry (MS): 0.3 µL samples were spotted onto a stainless-steel target and combined with 0.7 µL sample matrix (50% acetonitrile, 0.1% formic acid, 33% saturated solution of α-cyano-hydroxycinnamic acid). MALDI-ToF and MALDI-ToF/ToF experiments were performed on an Applied Biosystems 4800 Plus MALDI TOF/TOF Analyzer in positive ion reflector modes. The presence of DSPC was evaluated by MS/MS where the precursor mass was set to the mass of DSPC at m/z 790.6. In the resulting MS/MS spectrum, the presence of the phosphocholine headgroup at m/z 184.07 was considered diagnostic for the presence of DSPC. The MS was calibrated externally using the Bruker Daltonics Peptide Calibration Mixture and the MS/MS calibrated externally using the fragment ions of the angiotensin I M+H ion (m/z 1298.68). A focus mass of m/z 1600 was used for the MS acquisitions. For MS/MS 1000 kV was used for the collision energy.

Viral Transduction Assays: Antiviral efficacy against Dengue virus (DENV) was assessed using the micro Foci Reduction Test (FRT) that was adapted to a 96-well plate format (mFRT) as described earlier (14, 15). Briefly, 96-well plates in blocking buffer for 1 hr at 37C. Anti-envelop (4G2) (ATCC, HB114) and anti-PrM (2H2) was added to fixed cells in blocking buffer and incubated for 1 hr at 37C. Cells were washed and anti-mouse conjugated to HRP (Kirkegaard and Perry Laboratories, KPL) was added and incubated for 1 hour. Cells were washed and TrueBlue HRP substrate (KPL) was added, and the resulting blue viral foci were counted in Elispot machine.

Zika (ZIKV H/PF 2013) virus assay was performed using a plaque reduction test. Vero cells were seeded at a density of 5×10^4 cells per well in a 24 well plate, and infected with AEG12-treated ZIKV H/PF 2013, untreated ZIKV H/PF 2013m, or PBS as described above. Infected plates were washed with PBS then overlaid with 1 ml of 1% methyl cellulose (Sigma M0512) in 1X Opti MEM + 2% FBS, P/S, L-glutamine and incubated for 4 days at 37C, 5% CO_2. The methyl cellulose layer was subsequently removed, and cells were fixed with 2% crystal violet in 10% ethanol for one hour. Crystal violet was washed away with copious amount of tap water. Plates were dried and plaque numbers were counted.

Coronavirus (HCoV 229e) and Zika (ZIKV Paraiba 2015) virus plaque assays were performed by seeding MRC-5 or Vero E6 cells (respectively) to achieve 80-90% confluency on the day of transduction in 6-well plates. Cells were infected with AEG12-treated viruses along with the appropriate positive and negative controls as described above, and incubated for 1 hour. Vero cells were subsequently washed and overlaid with methylcellulose, incubated for 6 days at 37 °C, 5% CO_2, and the resulting plaques were visualized and counted as described above. For MRC-5 cells, coronavirus was diluted in EMEM+10% FBS and plates were incubated at 35°C, 5% CO_2. MRC-5 cells were overlaid with 0.8% methylcellulose (Sigma M0512) in 1X EMEM + 5% FBS, P/S, L-glutamine and incubated for 6-7 days. Cells were fixed with 2% crystal violet in 10% formalin in PBS for 2 hours, washed with tap water and dried overnight prior to counting the number of plaques.

Transduction assays for AAV and lentivirus were carried out using the GFP-modified strains (AAV-DJ EF1a-eGFP-WPRE and Lentivirus CD511B-2-EF1a-copGFP respectively). Assays were performed using 6-well plates seeded with 5×10^4 cells per well. Viruses were treated with the test compounds for 30 minutes at 37 °C, 5% CO_2. 1.26 µl of lentivirus CD511B-2-EF1a-copGFP (1.59×10^9 TU/ml (MOI = 10) or 5µl of AAV-DJ EF1a-eGFP-WPRE (3.07×10^{12} GC/ml) was applied to each well. Expression of fluorescent moieties and the number of GFP-positive cells (% total) were determined by flow cytometry (BD LSRFortessa, at the NIEHS Flow Cytometry Center) 48 hours post transduction.

Viral Strains and Culture: Viral aliquots of Coronavirus strain (229E), Zika Virus (Paraiba 2015), recombinant Adeno-associated virus (AAV), and recombinant lentiviruses were provided by the NIEHS Viral Vector Core. Coronavirus 229E was purchased from the ATCC (VR-740) and propagated in MRC-5 cell line (ATCC, CCL-171). Zika virus (Paraiba 2015) was a generous gift from the laboratory of Dr. Steve

Whitehead (NIH/NIAID) and was propagated in Vero E6 cell line (ATCC, CRL-1586). Coronavirus and Zika virus titres were determined by performing plaque assays in their corresponding propagation cell lines (described in the preceding section). rAAV was generated as described previously(16) using plasmids AAV EF1a-eGFP-WPRE (Addgene Cat# 135428), pAAV2/DJ (generous gift from Dr. Mark Kay's laboratory) and pHelper (Cell Biolabs Inc., Part No. 340202).(17) All recombinant lentiviruses were packaged in HEK293T/17 cells (ATCC, CRL-11268) according to published protocols (18). Briefly, 293T cells were transiently transfected with pMD2G (Addgene Cat# 12259), psPAX2 (Addgene Cat# 12260) and a lentiviral transfer vector expressing EF1a-copGFP (System Biosciences, CD511B-1). Supernatant was collected 48 hours post transfection and concentrated by centrifugation at 50,000 g for 2 hours. Pellets were resuspended in PBS and used for infection. All titres were determined by performing quantitative ddPCR to measure the number of lentiviral particles that integrated into the host genome. DENV (1-4) and Zika virus (H/PF 2013) were kindly provided by the de Silva laboratory at the University of North Carolina-Chapel Hill.

Fig. S1

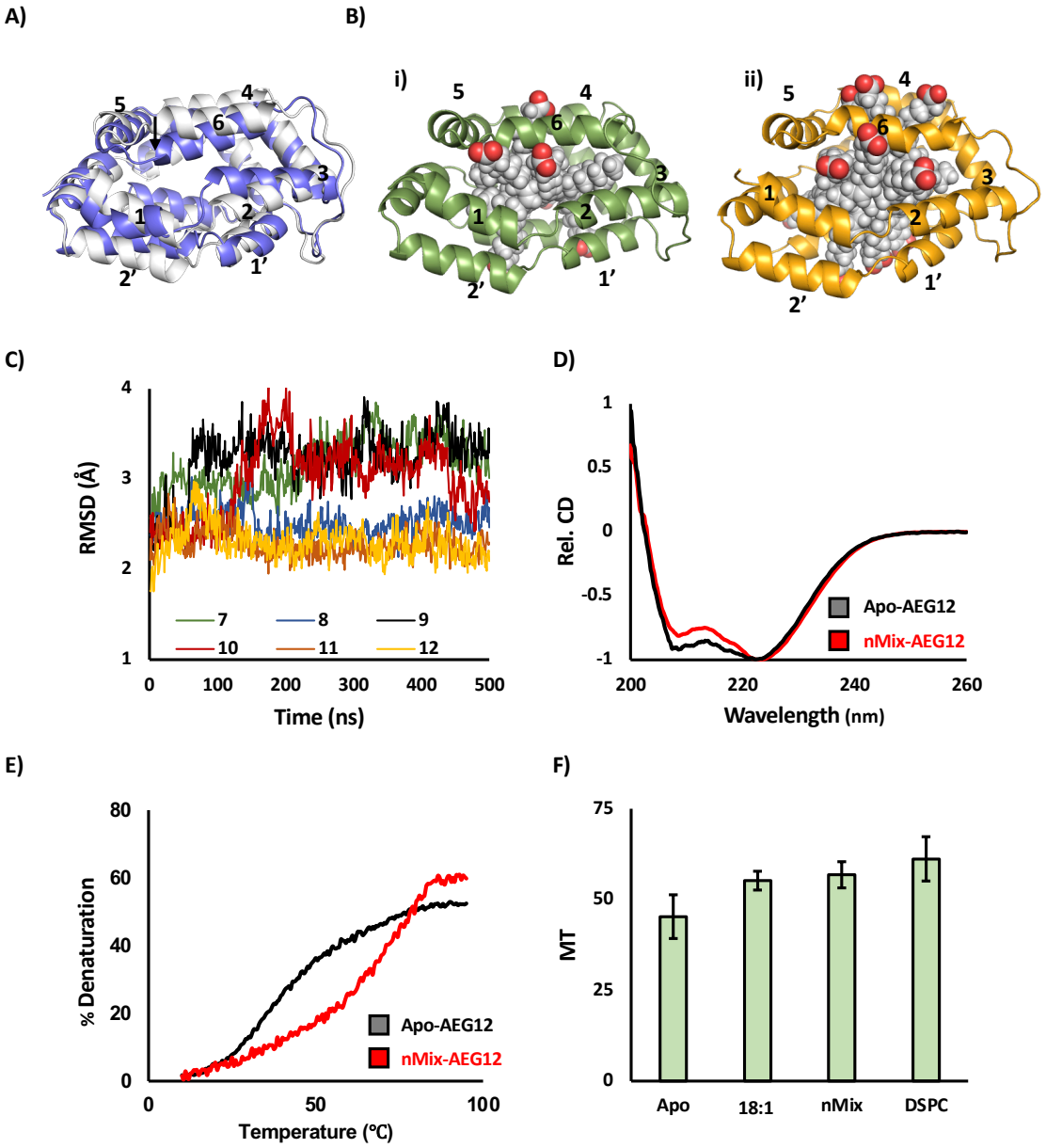
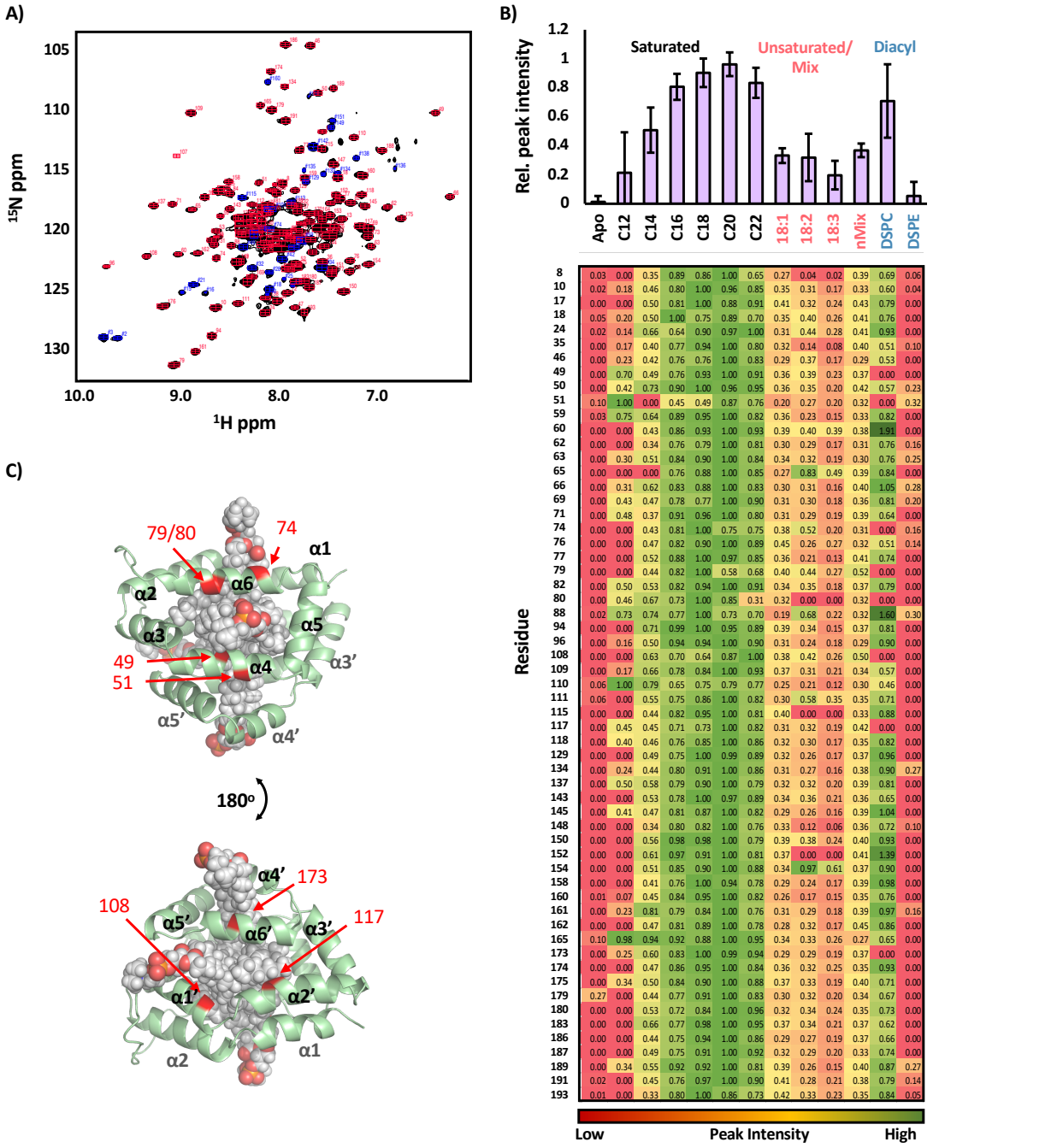


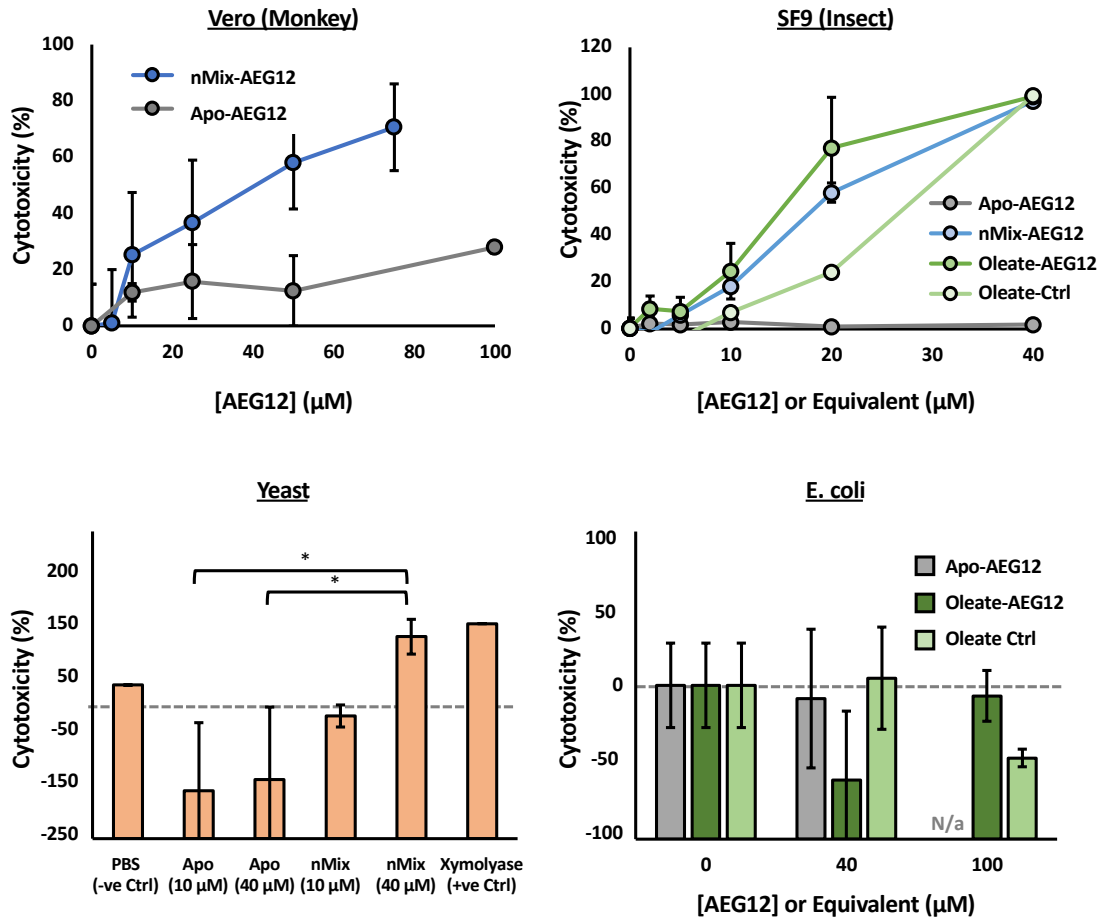
Fig. S1: Additional characterization of AEG12-ligand interactions. **A)** MD simulations of AEG12 bound to either 1 (Blue) or 4 (White) diacyl chain (DSPC) ligands (ligands omitted for image clarity). The former represents sub-stoichiometric binding and is accompanied by a collapse of helix 6 into the plane of the molecule (black arrow), along with minor rearrangements in the other helices to occupy the empty cavity **B)** MD simulations of AEG12 bound to 7 (i) or 12 (ii) oleate fatty-acid ligands showing minimal perturbation to the overall AEG12 scaffold. Backbone protein RMSD values obtained over the course of the 500 ns simulation window are shown in **(C)**. In no cases did the lipids disassociate from the protein during the studied time interval. **D)** Representative CD spectra of Apo- or nMix-AEG12 in Black and Red respectively. Spectra are normalized for protein concentrations and show minima at ~220 and 210 nm characteristic of a predominantly alpha-helical protein. Tracking the CD intensity at 220 nm vs temperature yields a thermal denaturation curve **(E)** from which melting temperatures (T_M), defined as the temperature at which 25% denaturation is observed, can be extracted. The T_M's for AEG12 loaded with various saturated and unsaturated fatty acid ligands at pH 7.5 are shown in Fig. 6A. Values obtained under acidic (pH 5.8) conditions are shown in **(F)**.

Fig. S2



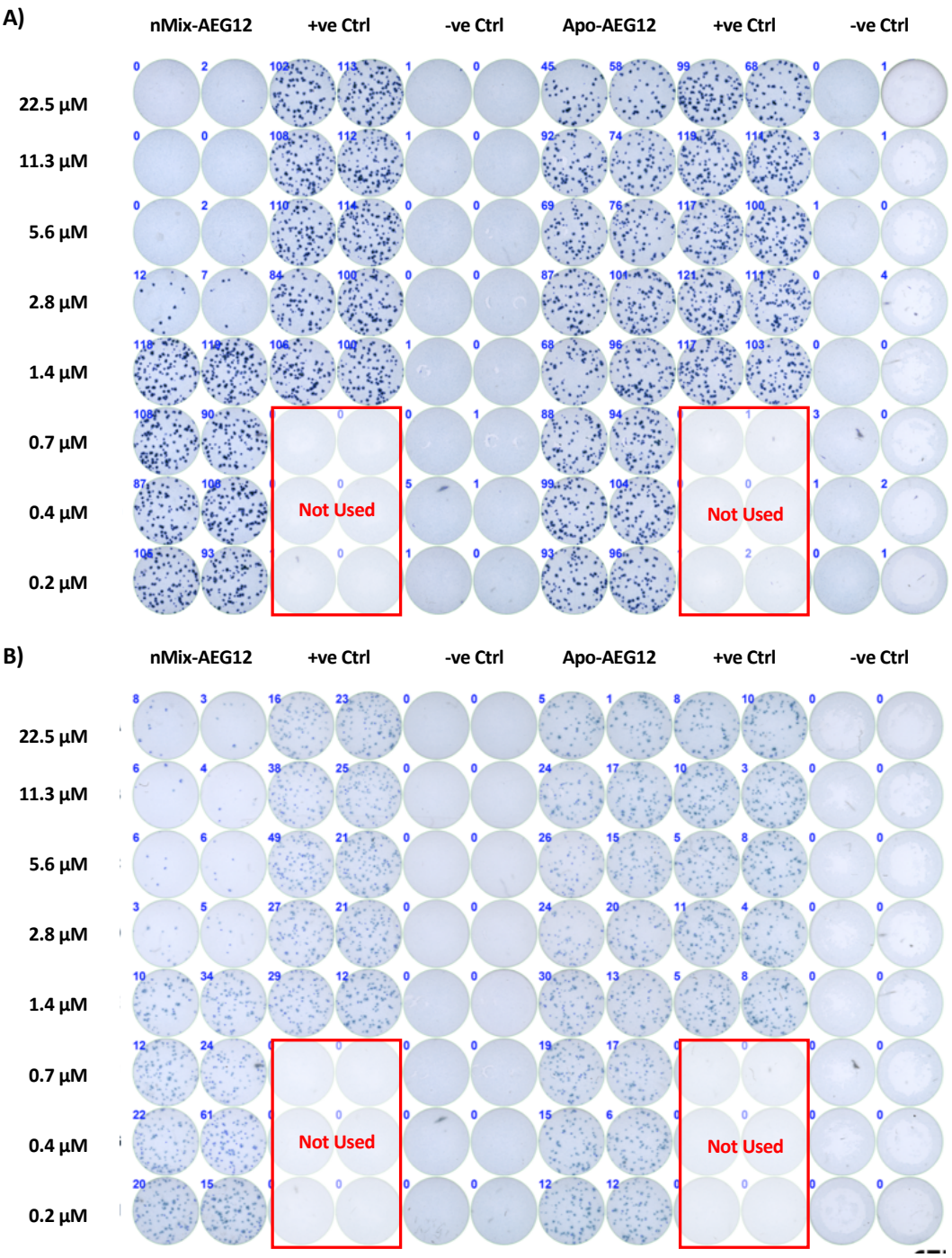
S2: Additional NMR characterization of AEG-12-ligand interactions. **A)** Backbone amide assignments of AEG12. Spectra obtained using perdeuterated ^{13}C-^{15}N-labeled AEG12. NMR assignments were deposited at the BioMagResBank #50434. **B)** Peak intensities for all identifiable AEG12 residues in the presence of various ligands. Values normalized against those obtained for AEG12 bound to saturated fatty acid ligands (ie: those falling within the Gaussian distribution). Mean values from Fig. 1 reproduced here for reference. Despite the overall high peak intensity values, a small number of peaks are noticeably missing from the DSPC-AEG12 spectrum. Plotting the location of these missing residues onto the structure of molecular dynamics simulation of AEG12 bound to a phospholipid cargo **(C)** shows that the missing residues (Red) are predominantly located near the mouth of the hydrophobic binding pocket. As such, their disappearance likely reflects reflect altered headgroup interactions rather than changes to global stability.

Fig. S3



S3: Cytolytic activity of AEG12 against a range of eukaryotic cells and *E. coli*. Note that negative values in the yeast and *E. coli* cytotoxicity data represent conditions in which treatment with AEG12 resulted in an increase in viable colonies relative to that of the negative control (PBS). Statistically significant differences (*) assessed using a student's t-test assuming equal variances. N/a denotes conditions which were not assessed in this study

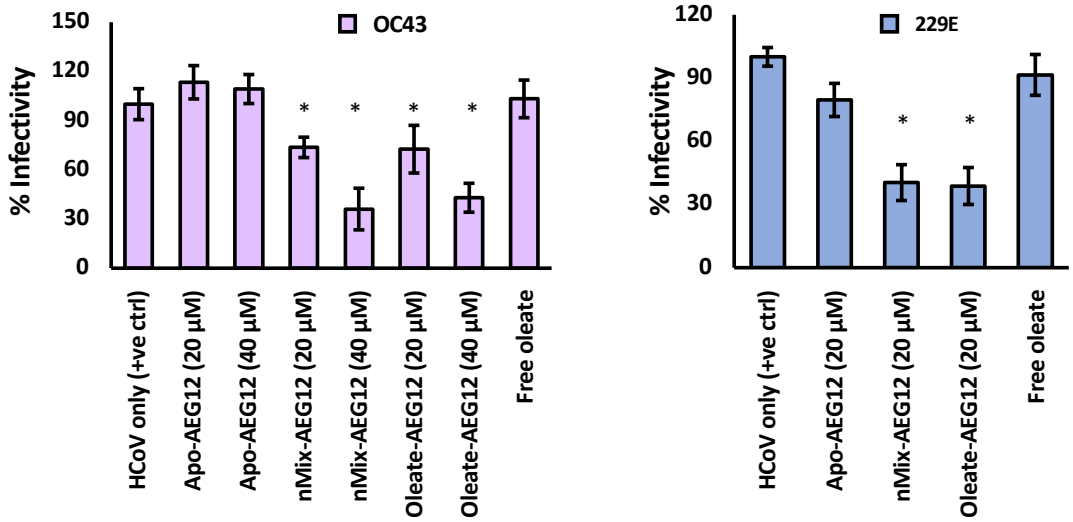
Fig. S4



S4: Antiviral Activity against DENV and ZIKV. Representative foci/plaque assay illustrating AEG12 antiviral activity against Dengue 1-4 (DENV) and Zikavirus ZIKV H/PF 2013 shown in **A)** and **B)** respectively. Viral foci (DENV) or plaques (ZIKV) indicated by blue stained dots. Positive (+ve Ctrl) and negative (-ve Ctrl) represents cells infected with untreated virus, or PBS respectively. Incubating the virus with varying concentrations of nMix- AEG12 significantly reduces infectivity as visualized by a decrease in the number of visible viral foci/plaques while Apo-AEG12 yields no reduction in infectivity relative to the control. Concentrations of AEG12 employed is shown on the left.

Fig. S5

A)

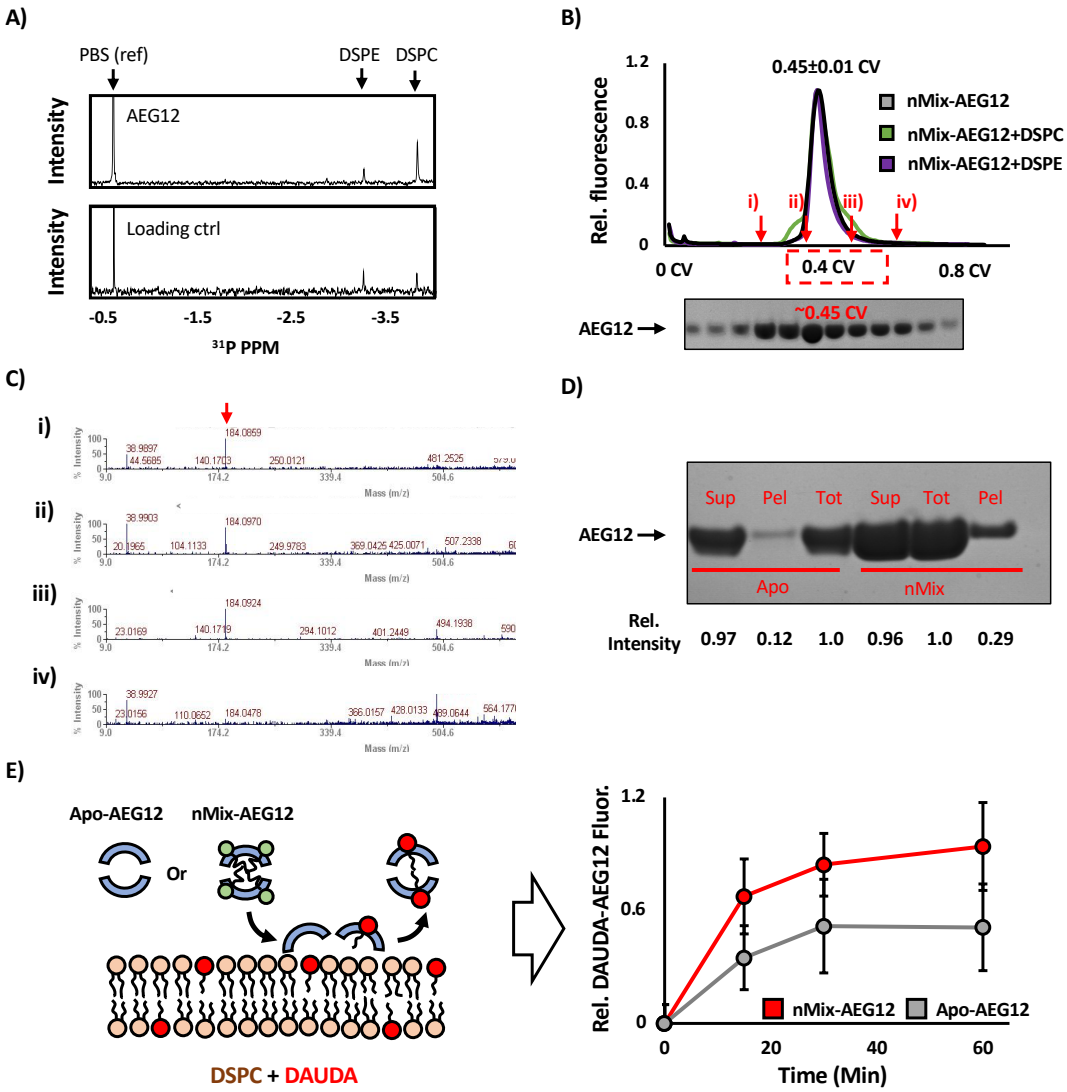


B)

	IC_{50} (µM) Apo-AEG12	IC_{50} (µM) nMix-AEG12	IC_{50} (µM) Oleate-AEG12
ZIKV (Paraiba 2015)	N/D	2.1±0.5	4.0±0.6
ZIKV (H/PF 2013)	N/D	3.7±0.5*	**N/a**
DENV	N/D	4.1±0.5*	**N/a**
HCoV 229	N/D	<20†	<20†
HCoV OC43	N/D	<40†	<40†
Lentivirus	N/D	48±3	36±2
AAV	N/D	**N/a**	N/D

S5: Additional Analysis of AEG12 antiviral activity. **A)** Antiviral activity of nMix-, Oleate-, and Apo-AEG12 against Human Coronavirus (HCoV) OC43 (Left) or 229e (Right). Conditions which differ significantly from the positive control, Apo-AEG12 treated samples, and free oleate controls as assessed using a student's t-test assuming equal variances are denoted with (*). **B)** Table of IC_{50} values for nMix-, Oleate-, and Apo-AEG12. Values obtained from two independent trials are denoted by (*). Values extrapolated from experiments in which AEG12 activity was assessed at a single concentration which yielded a ~50% reduction in infectivity as shown in A) are denoted with (†). N/D denotes conditions under which no antiviral activity was detected, while N/a denotes conditions which were not assessed in this study.

Fig. S6



S6: AEG12 Selectively binds phosphatidycholine (PC) lipids. **A)** ^{31}P-NMR spectra showing the phospholipid content of AEG12 subjected to the loading and annealing protocol in the presence of an equimolar mixture of DSPC and DSPE. ^{31}P-NMR spectra of the 1:1 DSPC:DSPE solution used to load AEG12 (loading control) is shown at the bottom for reference. A significant enrichment of DSPC is observed in the AEG12 sample despite the equimolar nature of the loading solution. **B)** Size-exclusion chromatography (SEC) elution profile for nMix-AEG12 in the absence or presence of DSPC lipid vesicles. Red rectangle illustrates the region analysed by SDS-PAGE shown below. Arrows indicate the fractions analyzed by MALDI-TOF MS shown in (C). **C)** MALDI-TOF spectra of AEG12+DSPC vesicles SEC elution fractions taken at various points in the elution profile. Phosphatidylcholine headgroups give a diagnostic peak at 184 m/z (red arrow), indicating the presence of DSPC in the AEG12 SEC peak. MALDI-TOF spectra of AEG12 incubated with DSPE vesicles yielded no detectable peaks. **D)** SDS-PAGE assay of Apo- or nMix- AEG12 incubated with DSPC vesicles. Vesicles were pelleted via centrifugation and gel samples taken for the supernatant (Sup) and resuspended pellet (Pel) fractions. Band intensities are normalized to the total (Tot) protein in both the soluble and vesicle-bound fractions prior to sedimentation. The resulting values are shown below and suggest that the majority of AEG12 remains in the soluble fraction. **E)** Formation of the DAUDA-AEG12 complex upon incubation with DAUDA-DSPC vesicles. DAUDA transfer to nMix- (Red) or Apo- (Grey) AEG12 monitored at 485 nm, corresponding to the DAUDA-AEG12 complex as shown in Fig. 4B. Fluorescence normalized against the equivalent volume of DAUSA-DSPC vesicles absent AEG12.

Table S1. Structural data for the X-ray crystal structure of AEG12.

Crystallographic data statistics		
Data Set		
Space Group	$P6_2$	$P6_2$
Unit cell	a, b = 112.46, c = 83.52	a, b = 113.01, c = 83.44
Resolution (Å)	2.1	1.95
# of observations	187,560	293,506
unique reflections	35,046	44,356
Rsym(%)(last shell)[1]	6.1 (38.5)	5.1 (79.2)
I/sI (last shell)	11.0 (2.5)	10.9 (2.3)
Mosaicity range	0.91 – 0.98	0.67 – 0.86
completeness(%) (last shell)	99.6 (97.1)	100.0 (100.0)
Refinement statistics		
Rcryst(%)[2]		17.7
Rfree(%)[3]		20.1
# of waters		549
Overall Mean B (Å) Protein Maltose Water		 25.5 11.3 32.7
r.m.s. deviation from ideal values		
bond length (Å)		0.003
bond angle (°)		0.581
dihedral angle (°)		15.72
Ramachandran Statistics[4]		
favored (>98%)		98.7%
allowed (>99.8%)		100%

1) Rsym = $\sum$ (| I_i - < I>|)/ $\sum(I_i)$ where I_i is the intensity of the ith observation and <I> is the mean intensity of the reflection.
2) Rcryst = $\sum$|| Fo| - | Fc ||/ $\sum$| Fo| calculated from working data set.
3) Rfree was calculated from 5% of data randomly chosen not to be included in refinement.
4) Ramachandran results were determined by MolProbity.

SI References

1. L. Shao, M. Devenport, H. Fujioka, A. Ghosh, M. Jacobs-Lorena, Identification and characterization of a novel peritrophic matrix protein, Ae-Aper50, and the microvillar membrane protein, AEG12, from the mosquito, Aedes aegypti. *Insect Biochem. Mol. Biol.* **35**, 947–959 (2005).
2. A. C. Y. Foo, *et al.*, Hydrophobic ligands influence the structure, stability, and processing of the major cockroach allergen Bla g 1. *Sci. Rep.* **9**, 1–12 (2019).
3. M. Cai, Y. Huang, R. Yang, R. Craigie, M. G. Clore, A simple and robust protocol for high-yield expression of perdeuterated proteins in Escherichia coli grown in shaker flasks. *J. Biomol. NMR* **66**, 85–91 (2017).
4. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr. Pt A* **276**, 307–326 (1997).
5. W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, HKL-3000: The integration of data reduction and structure solution - From diffraction images to an initial model in minutes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 859–866 (2006).
6. P. D. Adams, *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
7. A. J. McCoy, *et al.*, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
8. Y. Xu, *et al.*, Structure Based Substrate Specificity Analysis of Heparan Sulfate 6-O-Sulfotransferases. *ACS Chem. Biol.* **12**, 73–82 (2017).
9. G. A. Mueller, *et al.*, The novel structure of the cockroach allergen Bla g 1 has implications for allergenicity and exposure assessment. *J. Allergy Clin. Immunol.* **132** (2013).
10. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
11. V. B. Chen, *et al.*, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 12–21 (2010).
12. D. D. Case, *et al.*, AMBER 2016. University of California, San Francisco. (2016).
13. M. J. Frisch, *et al.*, Gaussian 09, Revision A.02 Gaussian, Inc., Wallingford, CT. (2016).
14. M. Montoya, *et al.*, Longitudinal Analysis of Antibody Cross-neutralization Following Zika Virus and Dengue Virus Infection in Asia and the Americas. *J. Infect. Dis.* **13**, 536–545 (2018).
15. M. . Collins, *et al.*, Cross-Neutralizing Antibodies Against Zika Virus from Dengue Virus Infection. *Emerg. Infect. Dis.* **23**, 773–781 (2017).
16. S. H. Chen, *et al.*, A Simple, Two-Step, Small-Scale Purification of Recombinant Adeno-Associated Viruses. *J. Virol. Methods* **281**, 113863 (2020).
17. D. Grimm, *et al.*, In Vitro and In Vivo Gene Therapy Vector Evolution via Multispecies Interbreeding and Retargeting of Adeno-Associated Viruses. *J. Virol.* **82**, 5887–5911 (2008).
18. P. Salmon, D. Trono, Production and titration of lentiviral vectors. *Curr. Protoc. Hum. Genet.* **54**, 12.10.1-12.10.24 (2007).