SCIENTIFIC REP🞮RTS

OPEN

Development of a S-adenosylmethionine analog that intrudes the RNA-cap binding site of Zika methyltransferase

Rinku Jain, Kyle V. Butler, Javier Coloma, Jian Jin 🄳 & Aneel K. Aggarwal

The Zika virus (ZIKV) has emerged as a major health hazard. We present here a high resolution structure (1.55 Å) of ZIKV NS5 methyltransferase bound to a novel S-adenosylmethionine (SAM) analog in which a 4-fluorophenyl moiety substitutes for the methyl group. We show that the 4-fluorophenyl moiety extends into a portion of the RNA binding tunnel that typically contains the adenosine 2′OH of the RNA-cap moiety. Together, the new SAM analog and the high-resolution crystal structure are a step towards the development of antivirals against ZIKV and other flaviviruses.

The Zika virus (ZIKV) belongs to the *Flavivirus* genus that includes mosquito-borne human pathogens such as dengue virus (DENV1-4), Murray Valley encephalitis virus (MVEV), West Nile virus (WNV), yellow fever virus (YFV), and Japanese encephalitis virus (JEV), among others[1]. ZIKV has emerged as a major health concern over the past year due to its proclivity to infect nerves *in utero* and in adults, leading to microcephaly in newborn infants and the Guillan-Barré syndrome in adults, respectively[2, 3]. Additional modes of transmission such as sexual transmission[4] and blood transfusions[5] threaten to hasten the spread of the virus. In addition to efforts to eradicate the *Aedes* mosquito vectors, there is an urgent need to develop effective vaccines and antiviral agents against ZIKV. One approach to the discovery of antivirals is to target enzymes that are essential for the replication and survival of the virus.

A methylated 5′-cap is a key feature of both flavivirus genomic RNA and host messenger RNA, required for RNA stability and translation[6]. In flaviviruses, the 5′-cap ($^{N7Me}GpppA_{2'OMe}$; where Me is a methyl group) is formed by the activities of two enzymes: NS3 that encodes a protease and helicase, and NS5 that encodes a methyltransferase (NS5-MTase) and an RNA dependent RNA polymerase. The NS5-MTase domain methylates both the N7 and 2′O atoms in a sequential reaction: $GpppA_{2'OH}$-RNA → $^{N7Me}GpppA_{2'OH}$-RNA or "cap-0" → $^{N7Me}GpppA_{2'OMe}$-RNA or "cap-1", using S-adenosylmethionine (SAM) as the methyl donor and generating S-adenosylhomocysteine (SAH) as the reaction byproduct[6].

Mutations in NS5-MTase that lead to defects in N7 methylation are lethal in flaviviruses[7–9]. Defects in 2′O methylation attenuate the virus and is a basis for vaccine development[10, 11]. Together, these features make ZIKV NS5-MTase an attractive target for the development of antivirals. We present here the design and characterization of a novel SAM analog (MS2042) with a 4-fluorophenyl moiety. We determine the thermodynamics parameters of binding of MS2042 to ZIKV NS5-MTase and elucidate a high-resolution (1.55 Å) structure of MS2042 bound to ZIKV NS5-MTase (NS5-MTase$_{MS2042}$). We show that the 4-fluorophenyl moiety extends into a portion of the RNA binding tunnel that typically contains the adenosine 2′OH of the cap-0 structure ($^{N7Me}GpppA_{2'OH}$-RNA). Together, the new SAM analog and the high-resolution crystal structure provide a path for further drug development.

Results

Design and synthesis of MS2042. We designed MS2042 on the basis of recent high resolution structures of the ZIKV NS5-MTase bound to SAM alone (NS5-MTase$_{SAM}$), or to both SAM and the 5′cap mimic 7-methyl guanosine diphosphate (7-MeGpp, NS5-MTase$_{SAM,7-MeGpp}$)[12]. We postulated that a 4-fluorophenyl group attached to the Cε atom of the methionine portion of SAM would have the capacity to bind a portion of the RNA binding tunnel that spans the SAM and 7-MeGpp binding sites. Supplementary Fig. 1a shows the scheme for the synthesis of MS2042.

Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, New York, USA. Correspondence and requests for materials should be addressed to A.K.A. (email: aneel. aggarwal@mssm.edu)
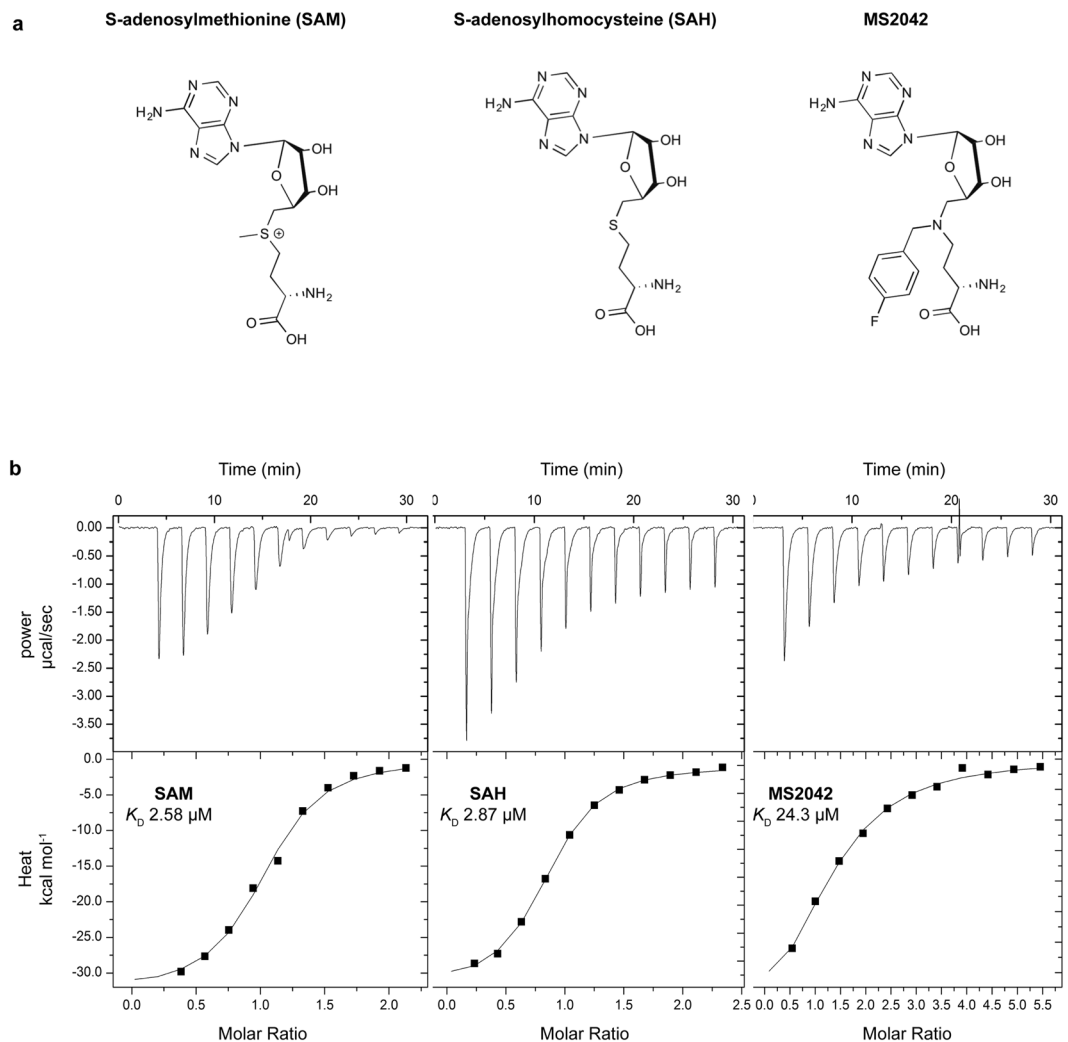
1

Figure 1. (**a**) Chemical structures and (**b**) Raw isothermal titration calorimetry data and binding isotherms for SAM (left), SAH (middle), and MS2042 (right) binding to ZIKV NS5-MTase.

Ligand	K_D (μM)	ΔG (Kcal mol^{-1})	ΔH (Kcal mol^{-1})	$-T\Delta S$ (Kcal mol^{-1})	K_D relative to that of SAM
SAM	2.58	−7.70	−32.6	24.9	—
SAH	2.87	−7.60	−39.8	32.2	1.11
MS2042	24.3	−6.30	−20.3	14.0	9.42

Table 1. Thermodynamic binding parameters of SAM, SAH, and MS2042 to ZIKV NS5-MTase.

Thermodynamic parameters of MS2042 binding to ZIKV NS5-MTase. A challenge in determining the binding of MS2042 to ZIKV NS5-MTase is in the preparation of the apo protein. This is because ZIKV NS5-MTase (residues 1–266) purifies from *E. coli* with SAM/SAH bound at the active site (NS5-MTase$_{SAM}$). To remove the bound SAM/SAH, we first denatured NS5-MTase$_{SAM}$ with 8 M urea and then refolded and purified the protein by dialysis and size exclusion chromatography (Supplementary Fig. 1b). We used isothermal titration calorimetry (ITC) to determine and compare the thermodynamic parameters of SAM, SAH, and MS2042 binding to refolded ZIKV NS5-MTase (Fig. 1). The enzyme binds SAM and SAH with similar affinities and equilibrium dissociation constant K_D values of 2.58 μM and 2.87 μM (Table 1), respectively. By comparison, the enzyme binds MS2042 with a K_D of 24.3 μM. Thus, the ZIKV enzyme can effectively bind MS2042 with its extra 4-fluorophenyl group, with an affinity that is ~9.4 lower than SAM (Table 1). To elucidate the structural determinants of MS2042 binding and how it compares to that of SAM, we determined the structure of MS2042 bound to ZIKV NS5-MTase (NS5-MTase$_{MS2042}$).

Overall structure of ZIKV NS5-MTase bound to MS2042. For crystallization, NS5-MTase and MS2042 were mixed in a 1:10 molar ratio. NS5-MTase$_{MS2042}$ crystallized from solution containing PEG 4000,

2-propanol, sodium citrate tribasic and glycerol. Crystals belong to space group P2(1) with unit cell dimensions of a = 38.99 Å, b = 111.23 Å, c = 77.40 Å, and β = 93.9° (Supplementary Table 1). The structure of NS5-MTase$_{MS2042}$ was determined by molecular replacement using the protein chain of NS5-MTase$_{SAM,7\text{-}MeGpp}$ (PDB id: 5KQS) as the starting model, and refined to 1.55 Å resolution with R$_{work}$ and R$_{free}$ values of 17.4% and 19.9%, respectively (Supplementary Table 1). The asymmetric unit contains two molecules of NS5-MTase (molecules A and B) related by two fold non-crystallographic symmetry (Supplementary Fig. 2). The two molecules are nearly identical and superimpose with an r.m.s deviation of 0.12 Å for 254 main chain Cα atoms. Changes are limited mainly to the side chain conformations of solvent exposed loop regions, including amino acids 108–110 (adjacent to the SAM binding cavity) and amino acids 253–255 that mediate contacts between the two molecules in the asymmetric unit. Overall, the core of ZIKV NS5-MTase$_{MS2042}$ adopts a Rossmann fold (residues 54–223) with seven β-strands (β1-β7) and four α-helices (αX, αA, αD, αE). The core is flanked by α-helices (A1-A4) and β-strands (B1-B2) that form the N- and C-terminal extensions and cradle the core (Fig. 2a). Amino acids belonging to the catalytic tetrad (Lys$_{61}$-Asp$_{146}$-Lys$_{182}$-Glu$_{218}$; a feature of 2′O methyltransferases) adopt the same conformation as in ZIKV NS5-MTase$_{SAM}$ and NS5-MTase$_{SAM,7\text{-}MeGpp}$ structures. Overall, the NS5-MTase$_{MS2042}$ structure is very similar to the NS5-MTase$_{SAM}$ and NS5-MTase$_{SAM,7\text{-}MeGpp}$ structures with r.m.s deviations of 0.17 Å (247 Cα atoms) and 0.18 Å (254 Cα atoms) respectively, indicative of little or no change in the protein conformation in accommodating the extra 4-fluorophenyl group of MS2042 (Supplementary Fig. 2b)[12].

Mode of MS2042 binding.　The electron density for MS2042 is clear in both molecules of the asymmetric unit (Fig. 2c). The SAM portion of MS2042 overlays almost exactly with the SAM in the NS5-MTase$_{SAM}$ and NS5-MTase$_{SAM,7\text{-}MeGpp}$ structures (Supplementary Fig. 2c), wherein the adenine ring is sandwiched between the side chains of Lys105 (partially disordered for chain B) and Ile147, and the N1 and N6 atoms establish hydrogen bonds with the backbone amide of Val132 and the side chain of Asp131 (Fig. 2d). In addition, the ribose sugar participates in solvent mediated interactions with Gly106, Glu111 and Thr104, while the amino and carboxyl groups of the methionine portion are held in place by hydrogen bonds with the side chains of Ser56 and Asp146, and the backbone amides of Gly86 and Trp87. A minor difference between molecules A and B in the NS5-MTase$_{MS2042}$ structure is that whereas the 2′OH of the MS2042 ribose in molecule A interacts with the side chain of His110, in molecule B His110 is pushed away into the solvent, in a conformation similar to that in the SAM bound structures (Supplementary Fig. 2c).

The 4-fluorophenyl group of MS2042 extends into a shallow cavity, adjacent to the SAM binding pocket (Fig 2b). The orientation of the phenyl ring differs slightly between the two molecules due to an ~11° rotation about the Nδ-Cε bond of SAM (Fig. 2c). There are no specific hydrogen bonds with any amino acids and the fluorine atom projects into the solvent. Most of the contacts between the 4-fluorophenyl group and the NS5-MTase are van der Waals in nature, wherein one edge of the phenyl ring (C2-C3) runs parallel to amino acids (146-Asp-Ile-Gly-Glu-149) in the NS5-MTase loop connecting β4 and αD (Fig. 3c) of the Rossmann core. Most importantly, the shallow cavity in the which 4-fluorophenyl group lies is the putative binding site of the base and the 2′OH group of the cap-0 adenosine (^{N7Me}GppA$_{2'OH}$-RNA) when RNA binds a flavivirus NS5 methyltransferase. Figure 3 shows the cap-0 moiety (^{N7Me}GppA$_{2'OH}$) modeled in our NS5-MTase$_{MS2042}$ structure, based on the structure of DENV3 NS5 harboring a cap-0 containing RNA substrate[13]. From our modeling, adenine base and the 2′O atom of the cap-0 structure sterically overlap with the phenyl ring of MS2042 and would be hindered in being methylated (Fig. 3a,b).

Discussion

There has been a significant effort to identify small molecule inhibitors for flavivirus MTases[14–20]. Interestingly, a DENV-selective SAM/SAH analog was identified, in which a chlorophenyl moiety is attached the adenine N6 atom of SAM/SAH[9, 16]. The chlorophenyl moiety extends into a small hydrophobic cavity located near the adenine of SAM/SAH. Since this cavity is mutually exclusive to the cavity that binds the 2′OH of the cap-0 structure, SAM/SAH analogs with dual substituents (one on the N6 atom of adenine and the other on Cε atom of the methionine portion of SAM/SAH) may provide additional selectivity for flavivirus MTases. In this context, the 4-fluorophenyl group on MS2042 also provides a basis for further chemical modifications. In particular, one edge of the 4-fluorophenyl ring (C2-C3) (Fig. 3c) runs parallel to a stretch of amino acids (146-Asp-Ile-Gly-Glu-149) in the NS5-MTase loop connecting β4 and αD. Based on our modeling, an exocyclic OH at the C2 position and/or NH at the C3 position would have the capacity to interact directly with amino acids on this β4-αD loop. Moreover, this loop is conserved in all flavivirus MTases but is different in sequence and structure in the human nucleoside-2′-O methyltransferase (CMTr1, PDB id 4N49)[21] and mRNA cap guanine-N7 methyltransferase (RNMT; PDB id 3EPP). The β4 -αD loop of human CMTr1 (364-Asp-Gly-Gly-Phe-367) and RNMT (122-Phe-Val-Cys-His-Tyr-126) is pushed up against the 4-fluorophenyl moiety of a modeled MS2042, relative to its location in ZIKV NS5-MTase (Fig. 3c–e and Supplementary Fig. 3). Based on our modeling, substituents on the C2, C3 and F4 positions of the 4-fluorophenyl moiety would be prone to steric clashes with the β4-αD loop of CMTr1 (Fig. 3d) and RNMT (Fig. 3e). Accordingly, MS2042 derivatives with chemotypes on the 4-fluorophenyl moiety offer the prospect of enhancing ligand selectivity for flavivirus NS5-MTases, relative to the human mRNA cap MTases.

In conclusion, we present here a high resolution structure of ZIKV NS5-MTase bound to a novel SAM analog as a step towards the development of antivirals against ZIKV and other pathogenic flaviviruses.

Methods

Purification of Apo ZIKV NS5-MTase.　ZIKV NS5-MTase (1–266) from the H/PF/2013 strain was expressed and purified from *E. coli* strain LOBSTR (DE3) with a N-terminal His$_6$-SUMO tag. Cell pellets containing the recombinant protein were resuspended in buffer containing 50% B-PER (Thermo Scientific), 25 mM Tris,
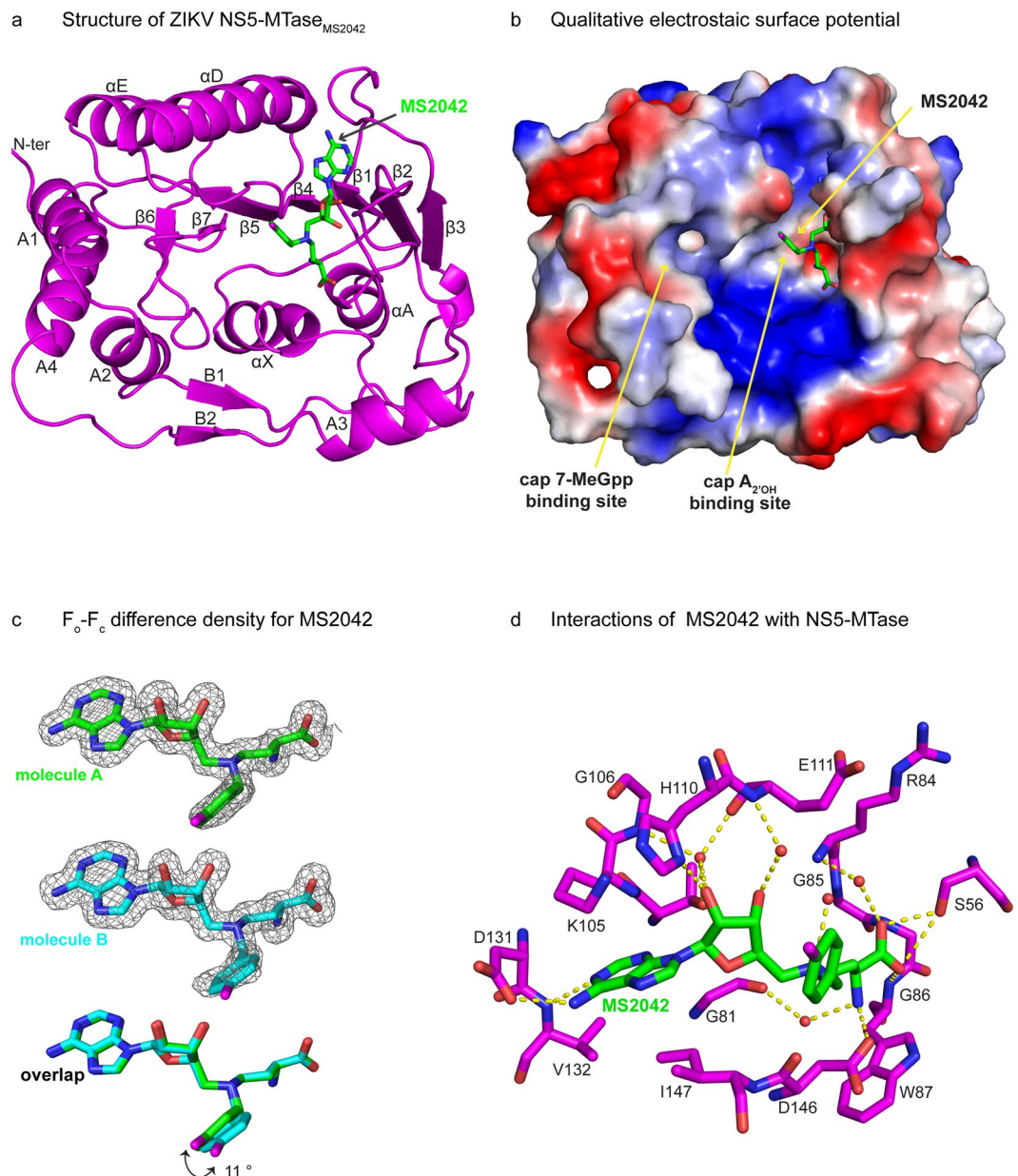
Figure 2. Structure of NS5-MTase$_{MS2042}$. (**a**) Overall structure of NS5-MTase$_{MS2042}$ with labeled secondary structures. (**b**) Qualitative electrostatic surface potential of NS5-MTase$_{MS2042}$ showing the shallow cavity accommodating the 4-fluorophenyl group of MS2042. The SAM portion of MS2042 is partially hidden from view (**c**) F_o-F_c difference density over MS2042 contoured at 2.5σ. Superimposition of MS2042 bound to molecules A and B reveals a slight difference in the orientation of the 4-fluorophenyl groups. (**d**) Interactions of MS2042 with the protein chain. Hydrogen bonds and solvent molecules are depicted as dashed lines and red spheres, respectively.

pH 8.0, 500 mM NaCl, 5% glycerol, and 5 mM 2-mercaptoethanol (BME). Cells were lysed by sonication and the filtered lysate was loaded on a 5 mL Ni-NTA column (Qiagen). Protein bound to the Ni-NTA column was eluted with buffer containing 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol, 5 mM BME and 250 mM imidazole. Eluted protein was dialyzed into buffer containing 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, and 5 mM BME. The His$_6$-SUMO tag was cleaved with Ulp protease and the protein re-loaded on the Ni-NTA column to remove the cleaved His$_6$-SUMO tag and any uncleaved protein. To remove SAM/SAH bound to the purified protein, the protein was denatured by adding urea to a final concentration of 8 M and overnight dialysis against buffer containing 50 mM Tris, 500 mM NaCl and 8 M urea. Refolding was initiated by dialyzing overnight into buffer containing 25 mM HEPES, pH 7.0, 250 mM NaCl, 2 mM TCEP, 5% glycerol, and 1 M urea. The refolded protein was purified further by size exclusion chromatography on a Superdex 75 16/600 GL column (GE Healthcare Life Sciences). Before crystallization, the protein was concentrated to 3.5 mg/ml in buffer containing 25 mM HEPES, pH 7.0, 250 mM NaCl, 5% glycerol and 2 mM TCEP.
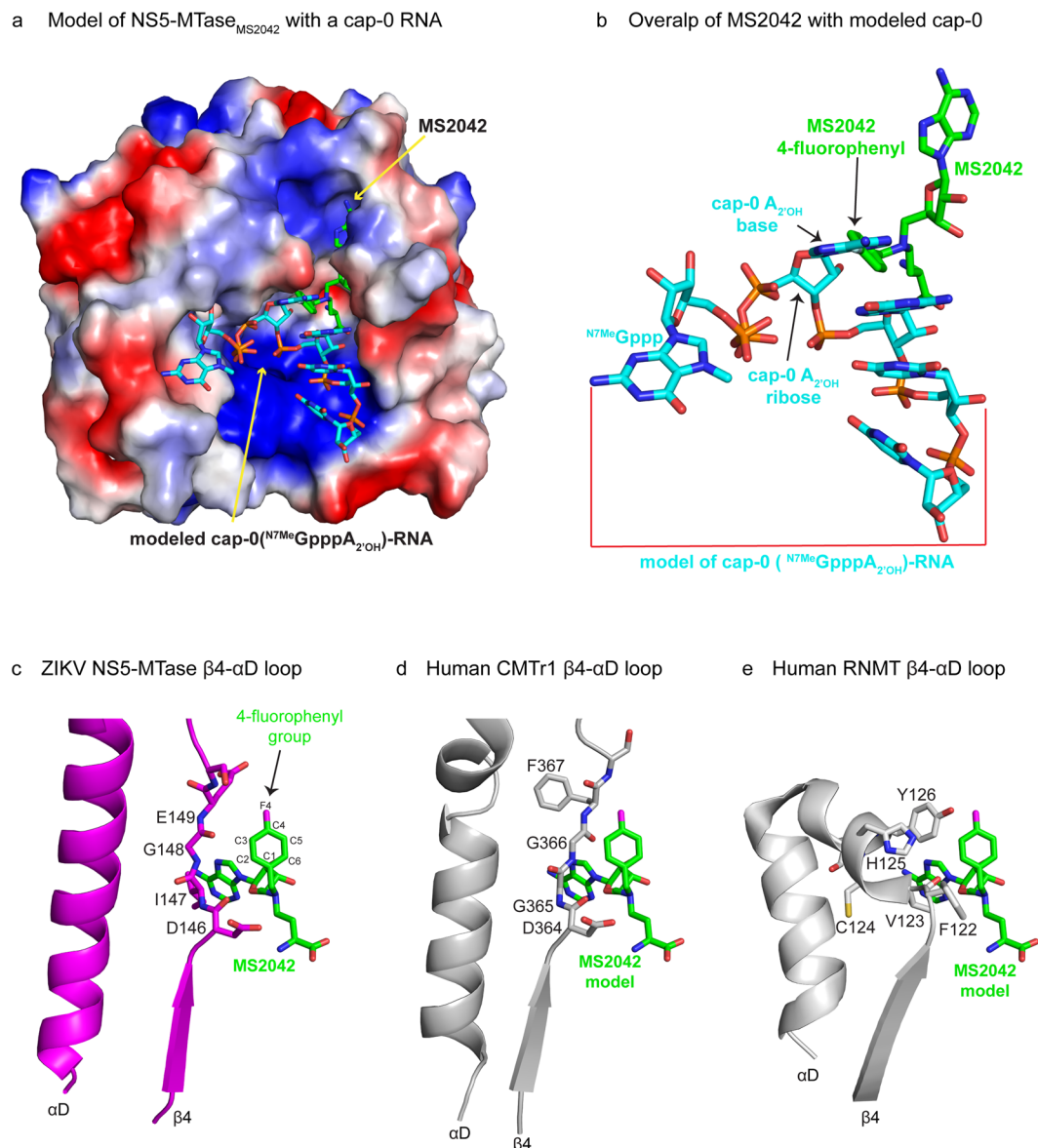
Figure 3. Model of ZIKV NS5-MTase$_{MS2042}$ with cap-0 RNA and comparison with the human mRNA cap methyltransferases CMTr1 and RNMT. (**a**) Qualitative electrostatic surface potential of ZIKV NS5-MTase$_{MS2042}$ with a cap-0(^{N7Me}GpppA$_{2'OH}$) RNA model derived from the structure of DENV3 NS5 (5DTO). (**b**) Magnified view of the model showing steric overlap of the cap-0 adenine and its 2′O atom with the 4-fluorophenyl group of MS2042. Comparison of loop β4-αD in (**c**) ZIKV NS5-MTase, (**d**) human nucleoside-2′-O methyltransferase (CMTr1, PDB id 4N49), and (**e**) human mRNA cap guanine-N7 methyltransferase (RNMT, PDB id 3EPP).

Synthesis of MS2042.

General Chemistry Procedures. All reagents were purchased from Sigma-Aldrich or Fisher Scientific. Solvents used in reactions were dried before use unless otherwise noted. Analytical HPLC – Method A: data were acquired using an Agilent 6110 series system with UV detection at 254 nm. Column: Agilent Eclipse Plus 4.6 mm × 50 mm, 1.8 um C$_{18}$ column. Solvents: A: 0.1% acetic acid in water; B: 0.1% acetic acid in methanol. Gradient: 10% to 100% B over 5.0 min, followed by 100% B for 2 min, at 1.0 mL/min. Method B: data were acquired using an Agilent Zorbax 300SC-C18 (5 μm) column with UV detection at 254 nm on an Agilent 1200 Series LC/MSD TOF machine. Solvents: A: 0.1% acetic acid in water; B: 0.1% acetic acid in methanol. Gradient: 1% B for one minute, 1 to 100% B over 3.0 min, followed by 100% B for 4 min, at 1.0 mL/min. Low resolution mass spectrometry (LRMS) data were acquired in positive ion mode on an Agilent 6110 single quadrupole mass spectrometer with electrospray ionization. Nuclear magnetic resonance (NMR) spectra were recorded on a Varian Mercury spectrometer at 400 MHz for proton (^{1}H NMR) and 100 MHz for carbon (^{13}C NMR). In other cases, NMR were recorded on a Bruker DRX spectrometer at 600 MHz for proton (^{1}H NMR) and 150 MHz for carbon (^{13}C NMR). Chemical shifts are reported in ppm (δ). Preparative HPLC was performed on an Agilent Prep 1200 series with UV detector set to 254 nm, using a Phenomenex Luna 75 mm × 30 mm, 5 um C$_{18}$ column, with a flow rate of 30 mL/min. High resolution mass spectrometry (HRMS) data was acquired on an Agilent 1200 Series

LC/MSD TOF. Medium pressure liquid chromatography (MPLC) was performed on a Combiflash Isco machine. Final compounds had >95% purity as judged by analytical HPLC unless noted.

Synthesis scheme of MS2042 ((S)-2-amino-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(4-fluorobenzyl)amino)butanoic acid). 6-Amino-N6,N6-bis(tert-butoxycarbonyl)-9-[5,6-dideoxy-6-(ethoxysulfonyl)-2,3-O- isopropylidene-β-D-ribo-hex-5-enofuranosyl]-9H-purine was prepared as described[22]. 6-Amino-N6,N6-bis(tert-butoxycarbonyl)-9-[5,6-dideoxy-6-(ethoxysulfonyl)-2,3-O-isopropylidene-β-D-ribo-hex-5-enofuranosyl]-9H-purine (350 mg, 0.86 mmol) and N-α-Boc-L-2,4-diaminobutyric acid tert-butyl ester hydrochloride (249 mg, 0.8 mmol) were stirred in 10 mL MeOH for 3 hours and then evaporated to dryness. The solid was dissolved in 20 mL of 1,2-dichloroethane, treated with acetic acid (60 μL) and sodium triacetoxy borohydride (212 mg, 1 mmol). This was stirred for 24 hours than diluted with saturated aqueous sodium bicarbonate. The product was extracted with dichloromethane, washed with brine, and purified on silica gel (gradient of 0 to 50% ethyl acetate in hexane). HRMS revealed the mono-alkylated product. HRMS (ESI-TOF): $[M + H]^+$ Expected: 764.4194; Found: 764.4209. The resulting solid (60 mg) was dissolved in MeOH (2 mL) and treated with: 4-fluoro-benzaldehyde (12.5 μL, 0.12 mmol), acetic acid (7.2 μL, 0.12 mmol), and sodium cyanoborohydride (8 mg, 0.12 mmol). This was stirred for 24 hours and then treated with saturated aqueous sodium bicarbonate. The product was extracted with dichloromethane, and the combined organic extracts were dried to a residue. This residue was treated with acetonitrile (1.5 mL), water (0.5 mL) and trifluoroacetic acid (1.5 mL) and stirred for 2 hours. Purification by HPLC gave the product as a trifluoroacetic acid salt. Yield: 15 mg; (0.025 mmol, 3% over three steps) HRMS (ESI-TOF): $[M + H]^+$ Expected: 476.2058; Found: 476.2066. ¹H NMR (400 MHz, MeOD, Supplementary Fig. 4) δ 8.38 (s, 1H), 8.31 (s, 1H), 7.50 (m, 2H), 7.11 (m, 2H), 6.12 (s, 1H), 4.63 (d, 1H), 4.46 (m, 3H), 4.30 (d, 1H), 3.84 (s, 1H), 3.52 (m, 3H), 2.34 (s, 1H), 2.13 (s, 1H). 13C NMR (100 MHz, CDCl₃) δ 170.76, 164.18, 162.53, 161.13, 152.41, 148.27, 142.27, 132.91, 126.29, 119.57, 115.64, 115.49, 90.80, 78.86, 73.22, 72.15, 56.82, 54.38, 52.11, 51.67, 25.04.

Isothermal titration microcalorimetry. Experiments were performed with MicroCal iTC$_{200}$ (GE Healthcare). Ligand SAM/SAH/MS2042 was loaded in the syringe and titrated into 60 μM of Apo ZIKV NS5-MTase in the cell. Care was taken to ensure buffer match for the ligand and ZIKV NS5-MTase to eliminate heat from buffer mismatch. The initial concentration used was 600 μM for SAM and SAH and 1 mM for MS2042. Titrations were performed at 25 °C with the standard 10 μcal/s reference power. Data were integrated, analyzed using the MicroCal iTC$_{200}$ Analysis Origin software provided with the instrument. All three ligands displayed 1:1 binding (within model error) and sigmoidal binding isotherms, suggestive of exothermic binding. The equilibrium dissociation constant[23] K_D for each ligand is shown in Table 1.

Crystallization and structure determination. To obtain co-crystals of ZIKV NS5-MTase and MS2042, a mixture was prepared by adding 10 fold molar excess of the ligand to Apo ZIKV NS5-MTase. The mixture was incubated on ice for ~15 minutes and initial screens were set up with the Oryx robot at 20 °C. The best looking crystals were obtained in conditions containing PEG 4000, 2-propanol, glycerol and sodium citrate trihydrate tribasic. For data collection, crystals were cryoprotected by quick dipping in perfluoropolyether cryo oil (Hampton Research) and flash-cooled in liquid nitrogen. Diffraction data were collected at the Advanced Photon Source (beamline 23-ID-D) under cryogenic conditions at a wavelength of 1.03324 Å, and indexed with HKL2000[24].

The ZIKV NS5-MTase$_{MS2042}$ structure was solved by molecular replacement using the Auto-Rickshaw web server (http://webapps.embl-hamburg.de/cgi-bin/Auto-Rick/arinitAR1.cgi)[25] and contains two molecules in the asymmetric unit. Model obtained from the Auto-Rickshaw pipeline was improved by iterative manual building and refinement with Coot[26] and Phenix[27] respectively. After the protein chain was built, significant difference electron density (2.5σ) was visible in the SAH/SAM binding site (Fig. 2c). The electron density was consistent with the presence of MS2042 ligand in both the chains of the asymmetric unit. This difference density was modeled as MS2042. The slightly lower occupancy of the fluorine atom and the slightly different orientation of the 4-fluorobenzyl moiety of MS2042 bound to the two chains is suggestive of conformational flexibility in binding. The structure of NS5-MTase$_{MS2042}$ was refined to 1.55 Å (R$_{free}$ of 19.9%; R$_{work}$ of 17.3%). The model has excellent stereochemistry as shown by MolProbity[28] with 99.5% of all residues in allowed regions of the Ramachandran plot and 0.5% in the disallowed regions. The residues in the disallowed regions include Asp256 from Chain A and Glu255 and Asp256 from chain B. These residues constitute a stretch of amino acids (253–258) that mediate interactions between the two molecules of the asymmetric unit and have disordered density.

Modeling. The model of ZIKV NS5-MTase$_{MS2042}$ bound to RNA was obtained by using the x-ray structure of full length DENV3 NS5 bound to an octameric RNA harboring a cap-0 (^{N7Me}Gppp) moiety at its 5′ end (PDB ID: 5DTO)[13]. The structures were superimposed via the NS5-MTase domains using PyMol (https://www.pymol.org/). ZIKV and DENV NS5-MTase share high sequence similarity and superimpose with a r.m.s. deviation of 0.46 Å for 210 Cα atoms.

MS2042 was modeled in the human mRNA nucleoside-2′-O methyltransferase (human CMTr1, 4N49) and mRNA cap guanine-N7 methyltransferase (human RNMT, 3EPP) by first superimposing the proteins via the Rossmann fold and then optimizing the alignment so good overlap was achieved between MS2042 bound to ZIKV NS5-MTase and SAM/sinefungin in the structure of human CMTr1/RNMT (Supplementary Fig. 3). Qualiative electrostaic surface potentials and figures were prepared using PyMol.

References
1. Petersen, L. R., Jamieson, D. J., Powers, A. M. & Honein, M. A. Zika Virus. *N Engl J Med* **374**, 1552–63, doi:10.1056/NEJMra1602113 (2016).

2. Fauci, A. S. & Morens, D. M. Zika Virus in the Americas–Yet Another Arbovirus Threat. *N Engl J Med* **374**, 601–4, doi:10.1056/NEJMp1600297 (2016).
3. Lazear, H. M. & Diamond, M. S. Zika Virus: New Clinical Syndromes and its Emergence in the Western Hemisphere. *J Virol* (2016).
4. D'Ortenzio, E. *et al.* Evidence of Sexual Transmission of Zika Virus. *N Engl J Med* **374**, 2195–8, doi:10.1056/NEJMc1604449 (2016).
5. Motta, I. J. *et al.* Evidence for Transmission of Zika Virus by Platelet Transfusion. *N Engl J Med* **375**, 1101–3, doi:10.1056/NEJMc1607262 (2016).
6. Dong, H. *et al.* Flavivirus RNA methylation. *J Gen Virol* **95**, 763–78, doi:10.1099/vir.0.062208-0 (2014).
7. Ray, D. *et al.* West Nile virus 5′-cap structure is formed by sequential guanine N-7 and ribose 2′-O methylations by nonstructural protein 5. *J Virol* **80**, 8362–70, doi:10.1128/JVI.00814-06 (2006).
8. Zhou, Y. *et al.* Structure and function of flavivirus NS5 methyltransferase. *J Virol* **81**, 3891–903, doi:10.1128/JVI.02704-06 (2007).
9. Dong, H. *et al.* Biochemical and genetic characterization of dengue virus methyltransferase. *Virology* **405**, 568–78, doi:10.1016/j.virol.2010.06.039 (2010).
10. Li, S. H. *et al.* Rational design of a flavivirus vaccine by abolishing viral RNA 2′-O methylation. *J Virol* **87**, 5812–9, doi:10.1128/JVI.02806-12 (2013).
11. Zust, R. *et al.* Rational design of a live attenuated dengue vaccine: 2′-o-methyltransferase mutants are highly attenuated and immunogenic in mice and macaques. *PLoS Pathog* **9**, e1003521, doi:10.1371/journal.ppat.1003521 (2013).
12. Coloma, J., Jain, R., Rajashankar, K. R., Garcia-Sastre, A. & Aggarwal, A. K. Structures of NS5 Methyltransferase from Zika Virus. *Cell Rep* **16**, 3097–102, doi:10.1016/j.celrep.2016.08.091 (2016).
13. Zhao, Y. *et al.* Molecular basis for specific viral RNA recognition and 2′-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5). *Proc Natl Acad Sci USA* **112**, 14834–9, doi:10.1073/pnas.1514978112 (2015).
14. Benmansour, F. *et al.* Discovery of novel dengue virus NS5 methyltransferase non-nucleoside inhibitors by fragment-based drug design. *Eur J Med Chem* **125**, 865–880, doi:10.1016/j.ejmech.2016.10.007 (2017).
15. Milani, M. *et al.* Flaviviral methyltransferase/RNA interaction: structural basis for enzyme inhibition. *Antiviral Res* **83**, 28–34, doi:10.1016/j.antiviral.2009.03.001 (2009).
16. Lim, S. P. *et al.* Small molecule inhibitors that selectively block dengue virus methyltransferase. *J Biol Chem* **286**, 6233–40, doi:10.1074/jbc.M110.179184 (2011).
17. Brecher, M. *et al.* Novel Broad Spectrum Inhibitors Targeting the Flavivirus Methyltransferase. *PLoS One* **10**, e0130062, doi:10.1371/journal.pone.0130062 (2015).
18. Coutard, B. *et al.* Assessment of Dengue virus helicase and methyltransferase as targets for fragment-based drug discovery. *Antiviral Res* **106**, 61–70, doi:10.1016/j.antiviral.2014.03.013 (2014).
19. Benarroch, D. *et al.* A structural basis for the inhibition of the NS5 dengue virus mRNA 2′-O-methyltransferase domain by ribavirin 5′-triphosphate. *J Biol Chem* **279**, 35638–43, doi:10.1074/jbc.M400460200 (2004).
20. Stahla-Beek, H. J. *et al.* Identification of a novel antiviral inhibitor of the flavivirus guanylyltransferase enzyme. *J Virol* **86**, 8730–9, doi:10.1128/JVI.00384-12 (2012).
21. Smietanski, M. *et al.* Structural analysis of human 2′-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat Commun* **5**, 3004, doi:10.1038/ncomms4004 (2014).
22. Shi, C. *et al.* Bisubstrate Inhibitors of Biotin Protein Ligase in Mycobacterium tuberculosis Resistant to Cyclonucleoside Formation. *ACS Medicinal Chemistry Letters* **4**, 1213–1217, doi:10.1021/ml400328a (2013).
23. Hulme, E. C. & Trevethick, M. A. Ligand binding assays at equilibrium: validation and interpretation. *Br J Pharmacol* **161**, 1219–37, doi:10.1111/j.1476-5381.2009.00604.x (2010).
24. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326, doi:10.1016/S0076-6879(97)76066-X (1997).
25. Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. Auto-rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr* **61**, 449–57, doi:10.1107/S0907444905001307 (2005).
26. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–32, doi:10.1107/S0907444904019158 (2004).
27. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–21, doi:10.1107/S0907444909052925 (2010).
28. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375–83, doi:10.1093/nar/gkm216 (2007).

Acknowledgements

Author Contributions

A.K.A., R.J., and J.C. conceived the project. R.J., A.K.A., K.V.B., and J.J. designed the experiments; R.J. performed protein expression, purification, crystallization, and ITC studies assisted by J.C.; K.V.B. performed the chemical synthesis of MS2042 under the guidance of J.J.; R.J. and A.K.A. wrote the paper. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-01756-7

Competing Interests: The authors declare that they have no competing interests.

Accession codes: Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 5ULP.