

A metal ion orients mRNA to ensure accurate 2'-O ribosyl methylation of the first nucleotide of the SARS-CoV-2 genome

Thiruselvam Viswanathan^{1,2,#}, Anurag Misra^{1,2,#}, Siu-Hong Chan³, Shan Qi^{1,2}, Nan Dai³, Shailee Arya¹, Luis Martinez-Sobrido⁴, Yogesh K. Gupta^{1,2,6,*}

¹ Greehey Children's Cancer Research Institute

² Department of Biochemistry and Structural Biology

University of Texas Health at San Antonio

8403 Floyd Curl Drive, San Antonio, TX 78229, USA

³ New England Biolabs, 240 County Road, Ipswich, MA, 01938, USA

⁴ Texas Biomedical Research Institute, San Antonio, TX, 78227, USA

⁶ Lead contact

*Corresponding author:

Y.K.G email: guptay@uthscsa.edu

These authors contributed equally

Abstract

The SARS-CoV-2 nsp16/nsp10 enzyme complex modifies the 2'-OH of the first transcribed nucleotide of the viral mRNA by covalently attaching a methyl group to it. The 2'-*O* methylation of the first nucleotide converts the status of mRNA cap from Cap-0 to Cap-1, and thus, helps the virus evade immune surveillance in the host cell. Here, we report two structures of nsp16/nsp10 representing pre- and post-release states of the RNA product (Cap-1). We observe overall widening of the enzyme upon product formation, and an inward twisting motion in the substrate binding region upon product release. These conformational changes reset the enzyme for the next round of catalysis. The structures also identify a unique binding mode and the importance of a divalent metal ion for 2'-*O* methylation. We also describe underlying structural basis for the perturbed enzymatic activity of a clinical variant of SARS-CoV-2, and a previous SARS-CoV outbreak strain.

Main

RNA viruses employ a diverse set of protein assemblies to enzymatically modify the 5'-end of their genomes. This process, termed RNA capping is essential for efficient a) production of viral proteins, b) protection of viral (v)RNA from host degradation, and c) subversion of the host innate immune responses – all of which enable viruses to thrive inside the host body¹. In coronaviruses (CoVs), the nonstructural protein 16 (nsp16), and its non-catalytic stimulator nsp10, assemble on the 5'-end of nascent mRNA to perform the last step of RNA cap modification – *S*-adenosyl-L-methionine (SAM)-dependent methylation of the 2'-OH on the first transcribed nucleotide base (N₁), usually an adenine. This converts the status of RNA from Cap-0 (^{me7}GpppA) to Cap-1 (^{me7}GpppAm)^{2,3}. Suppression of innate host antiviral response by Cap-1 through IFIT (interferon signaling-induced proteins with tricopeptide repeat) proteins⁴ is thought to be a consequence of diminished binding of IFIT to Cap-1⁵. Genetic ablation of nsp16 enzymatic activity also leads to induction of type I interferon (IFN) via the RNA sensor melanoma differentiation-associated protein 5 (MDA5)⁶. Emerging evidence suggests that COVID-19 patients have elevated innate immune responses, causing hypercytokinemia⁷. Thus, structural elucidation of different stages of Cap-1 formation and modification will further our understanding of the 5'-RNA biology of CoVs, and may inform the path to rational drug design.

We and others have previously resolved the structures of the SARS-CoV-2 nsp16/nsp10 enzyme complex in the presence of a Cap-0 analogue (^{me7}GpppA) and methyl donor SAM (*S*-adenosyl-L-methionine) (Fig. 1a)^{8,9}. Here we report two structures of the nsp16/nsp10 heterodimer complex in the presence of a cognate RNA product (Cap-1) that consists of N₁ and an adjoining N₂ base (^{me7}GpppAmU), and a byproduct of the methylation, SAH (*S*-adenosyl-L-homocysteine), resolved

to 2.3 and 2.5 Å, respectively (Fig. 1b-c, S. Table 1). These structures were solved by a molecular replacement method using the previously determined Cap-0 (^{m7}GpppA)/SAM structure (PDB ID: 6WKS, hereafter referred to as the “substrate structure”) as a search model⁸. Cap-1 RNA and SAH were unambiguously identified in the difference omit maps (S. Fig. 1 a-c). nsp16 adopts a canonical methyltransferase fold with a central β-sheet flanked by two α-helices on one side and three on the other, similar to the substrate Cap (Cap-0)-bound structure⁸ but with a notable difference as described below.

nsp16/nsp10 undergoes breathing motion during 2'-O methyl transfer

A structural comparison of nsp16/nsp10 in substrate, product (^{m7}GpppAmU and SAH), and byproduct (SAH only) bound structures revealed notable conformational changes in three complexes (Fig. 1). Most strikingly, we observed an overall expansion of the nsp16/nsp10 complex in the product structure as compared to the substrate-bound enzyme (~ 6.8 Å in one dimension and ~ 5.4 Å in the other) although the central β-sheet remains largely unperturbed (Fig. 1d-e, S. Fig. 1i). A positional shift in the Cap-1 analogue also occurs, so that the 2'-O-me group on the A₁ base now pushes the SAH outward, so that the sulfur atom points away from the 2'-O-me moiety, and the carboxy tail of SAH rotates 180° around C_β of the SAH (Fig. 1e, S. Fig. 1a-h, j). A comparison of product and byproduct structures reveals no major changes in protein conformations except for an inward shift in gate loop 2 (Fig. 1f). The position of SAH in the product and byproduct structures remains essentially the same, except for the carboxy tail, which rotates back in byproduct structure to assume the original orientation as in the substrate structure (Fig. 1b-d, S. Fig. 1g-h). These

changes suggest that the conformation of the SAH-bound enzyme represents a resetting for the next round of catalysis.

The surface area buried (BSA) at the nsp16/nsp10 interface is significantly smaller (843 \AA^2) in the product complex than the substrate structure (930 \AA^2) (Fig. 1d, g). This reduction is less pronounced (BSA= 878 \AA^2) in the byproduct structure (Fig. 1f). As a result, the heterodimeric interface in the product structure (and to some extent the byproduct structure) is widened by $\sim 1.5 \text{ \AA}$ at one end, $\sim 0.9 \text{ \AA}$ at the other, and $\sim 1.35 \text{ \AA}$ in the center (Fig. 1g-i). Thus, the much-relaxed heterodimeric interface in the product structure appears to be a result of overall widening of the enzyme, which is triggered by a single 2'-*O* methylation event. The enzyme appears to go into a “breathing motion” during catalysis, wherein the substrate, product, and byproduct bound states represent *fully closed*, *open*, and *partially open (product released/enzyme reset)* states, respectively (Fig. 1). The product RNA Cap, Cap-1 with adjoining uracil as N₂ base at 3'-end, is accommodated within a deep pocket constituted by one side of the central β -sheet and the two gate loops (gate loop 1 [amino acids 20-40] and gate loop 2 [amino acids 133-143]). The byproduct SAH in both structures binds similarly in a cavity at the C-terminal side of the parallel β -strands except for different orientations of their carboxy tails (S. Fig. 1a-b, g-h).

Metal dependency for 2'-*O*-methylation by SARS-CoV-2 and its clinical variants

We also observed unambiguous electron density in omit maps at the interface of gate loop 1, a loop between the $\beta 8$ and $\beta 9$ strands, and the phosphate moiety of uracil (U₂), the N₂ base downstream to the RNA Cap (Fig. 1e and S. Fig. 1a, c). The features of this density suggested a divalent metal ion coordinating with water molecules. Metal ions stabilize nucleic acid substrates

in addition to acting as catalytic agents in enzymatic reactions. In other positive single-stranded RNA viruses (e.g., 2'-*O* MTase such as dengue NS5), a magnesium ion stabilizes the RNA cap by coordinating with the inverted triphosphate moiety from the solvent-exposed side of the RNA cap (Fig 2f). A magnesium ion in nsp16 of the previous CoV outbreak strain (PDB ID: 2XYR) binds to a remote site constituted by T58 and S188 located at the opposite face of the catalytic pocket¹⁰. A direct binding of metals in the substrate/catalytic pocket and their role in 2'-*O* MTase activity of the CoVs nsp16, including SARS-CoV-2, has not been previously suspected.

Since our protein purification buffers and crystallization solutions contain magnesium and calcium salts, respectively, we fit these metals into the additional electron density, but could refine the product structure with better statistics when Mg²⁺ was modeled along with its coordinating water molecules (B factor = 52.5 Å² and 73.3 Å² for Mg²⁺ and Ca²⁺, respectively) in this density. As modeled, Mg²⁺ is further stabilized by direct interaction with the side chain of an invariant N198 and assumes a near-ideal octahedral geometry with coordinating water molecules (Fig. 2a-b, S. Fig. 1d-e, j). Consistently, the wild-type (WT), but not the N198A mutant of nsp16/nsp10 binds Mg²⁺ with high affinity (Fig. 2d). A direct interaction of protein to Mg²⁺ and its orientation in the Cap binding pocket is unique to SARS-CoV-2 (Fig. 2c). For example, in dengue NS5, Mg²⁺ is exposed to solvent and cross-links the phosphate groups of the RNA cap without ligating to the protein, whereas it directly binds to nsp16 (through N198) and the phosphate of the U₂ base (water-mediated) in SARS-CoV-2 (Fig. 2f).

Our previous work revealed the basis of target specificity of nsp16 for adenine nucleotide at the N₁ position⁸. We and others postulated that K170 (the second lysine of the KDKE catalytic tetrad)

acts as a general base to facilitate methyl transfer from SAM to 2'-OH of the A₁ nucleotide in SARS-CoV-2⁸ and other CoVs^{3,10}. The product structure allowed us to examine the roles of K170 and K46 (first lysine of the KDKE catalytic tetrad) in detail. The side chains of K170 and K46 form hydrogen bonds with the 2'-O of the product (i.e., methylated ribosyl of A₁) and phosphoryl oxygen of U₂ base, respectively (Fig. 2b, S. Fig. 1d-f). The phosphoryl oxygens of U₂ also interact with water molecules that coordinate with Mg²⁺. Mutation of K46 and N198 to alanine completely abolished Mg²⁺ binding and catalytic activity of nsp16/nsp10 (Fig. 2d-e). This is consistent with the network of side chains of K46, K170, and N198 that we identify here as important for catalysis. Such an arrangement correctly positions the RNA cap in the catalytic pocket to both ensure efficient 2'-O methylation of A₁ base and prevent unintended methylation of the adjoining U₂ base by restricting its movement or misalignment during catalysis of the A₁ base. The U₂ nucleotide is largely exposed to solvent and showed some deviation in the geometry (S. Fig. 1d).

Previously, we mapped the acquired mutations that have been identified in SARS-CoV-2 nsp16 on its structure and postulated their potential roles in RNA binding and/or catalysis⁸. Among these, the S33 residue in gate loop 1 is particularly noteworthy given the high occurrence of the S33R (20755:A>C) mutation in SARS-CoV-2 strain associated in the New York City outbreak¹¹, and its mutation to an asparagine (S33N) in a previous CoV outbreak strain (Fig. 2a, S. Fig. 1a). The side chain of S33 resides 6.3Å and 9.6Å away from Mg²⁺ and phosphoryl oxygens of the U₂ base in the product structure, respectively. The side chain of the arginine in S33R, as we modeled in Fig. 2c, may intrude into this pocket to reduce these distances by ~ 4.18Å, thus disrupting the coordination of Mg²⁺, which would in turn disorient the target base (A₁) in the catalytic pocket. A

shorter side chain of asparagine would be less intrusive and may provide additional contacts to a divalent metal ion, strengthen RNA binding, and therefore enhance 2'-*O*-methylation.

To further probe these models, we employed LC/MS to measure Cap-1 formation by nsp16/nsp10 WT and variant enzymes, and their dependence on various metal ions. Consistent to magnesium-binding affinity, where mutants N198A, K46A and S33R exhibited negligible Mg²⁺ binding (Fig. 2d), N198 (which directly interacts with magnesium), and K46 (which stabilizes the phosphate of U₂) when mutated to alanine, completely abolished the enzymatic activity of nsp16/nsp10 (Fig. 2e). Strikingly, the activity of S33R mutant decreased by ~ 80%, further validating our structural interpretation about this clinical variant (Fig. 2c-e). In contrast, the S33N mutation resulted in 30% increased activity, suggesting that the SARS-CoV nsp16, which has N at this position, has superior 2'-*O* methylation capability compared to SARS-CoV-2 nsp16 (Fig. 2e). Moreover, the SARS-CoV-2 nsp16 shows indistinguishable 2'-*O* methyltransferase activity in the presence of magnesium and manganese, but a 20% loss in the presence of calcium. The S33N mutant showed no preference for any of the three divalent ions tested (Mg²⁺, Ca²⁺, Mn²⁺). However, S33R showed comparable activity in the presence of Mg²⁺ and Mn²⁺ (although 80% less than the WT), but residual enzymatic activity in the presence of Ca²⁺ (Fig. 2e).

The unique role of a divalent metal ion in SARS-CoV-2 nsp16 appears to be architectural (Fig. 2f), yet it is essential for accurate and efficient 2'-*O* methylation of the first transcribed base of the SARS-CoV-2 genome. Thus, allowing the virus to evade host innate immune responses. Such reliance and preference for metals also suggests that an imbalance in cellular metal concentrations could differentially alter the RNA capping and thus, host innate immune response to infections by

various CoVs. In support of this concept, hypocalcemia is considered as a strong predictor of in-hospital COVID-19 deaths^{12,13}, and severely ill COVID-19 patients had significantly lower magnesium levels in whole blood¹⁴. One possibility is that the sub-optimal RNA capping by nsp16/nsp10 variants, together with altered levels of divalent metals, could trigger excessive immune response and cause hypercytokinemia in a subpopulation of COVID-19 positive patients. Future studies should determine the direct correlations between RNA capping, metal levels in the host cellular milieu, and innate immune response.

METHODS

Protein expression and purification

The nsp16 (NCBI reference sequence YP_009725311.1) and nsp10 (NCBI reference sequence: YP_0009725306.1) of the seafood market pneumonia SARS-CoV-2 isolate Wuhan-Hu-1 (NC_045512) were cloned into a duet vector, co-expressed in *E. coli*, and purified using the method we previously described⁸. Briefly, the clarified *E. coli* lysates were loaded on to a metal affinity column and proteins were eluted by running a concentration gradient of imidazole. The N-terminal His₆ tag from nsp16 was then proteolytically removed. The tag-free fractions were purified by successive passage through affinity, ion-exchange, and size-exclusion chromatography columns. The purified enzymes were concentrated to 5 mg/mL and used immediately for subsequent biochemical and/or crystallographic studies. We used the same method for all mutant nsp16/nsp10 enzymes reported in this study.

Crystallization, X-ray diffraction data collection and structure determination

The initial nsp16/nsp10 complex was grown by the sitting drop vapor diffusion in a crystallization solution 10% (v/v) of 2-propanol, 0.1 M MES/NaOH pH 6.0, 0.2 M calcium acetate. After 3-4 rounds of optimization by varying pH, precipitant, and salt concentrations, we grew larger crystals amenable to synchrotron radiation. We soaked these crystals with an RNA analog representative of the Cap-1 structure (^{me7}GpppA_(2'-O-me)U). The crystals were cryo-protected by serial soaks in a solution containing the original mother liquor and increasing concentrations (0 to 20% v/v) of ethylene glycol, and then flash-frozen in liquid nitrogen. Crystals of the nsp16/nsp10/Cap-1/SAH and nsp16/nsp10/SAH complexes diffracted X-rays to 2.3 and 2.5 Å resolution with synchrotron radiation, respectively (S. Table 1). Both crystals belong to the space group P3₁21 with similar

unit cell dimensions $a=b=184\text{\AA}$, $c = 57\text{\AA}$, $\alpha = \beta = 90^\circ$, and $\gamma = 120^\circ$, and with one nsp16/nsp10 heterodimer per asymmetric unit. All data (measured at wavelength 1.07 \AA) were indexed, integrated, and scaled using XDS, aimless, and various ccp4 suite programs integrated into the RAPD pipeline at the NECAT 24ID beamline¹⁵. The structure was solved by molecular replacement using a Cap-0 ternary complex of nsp16/nsp10/SAM (PDB ID: 6WKS)⁸ structure as a template in Phaser¹⁶. The resulting maps indicated unambiguous electron densities for RNA Cap-1 and SAH. Ligand topologies and geometrical restraints were generated using PRODRG (<http://prodrgr1.dyndns.org>), GRADE (<http://grade.globalphasing.org>), and eLBOW (Phenix)¹⁶ programs. We iteratively rebuilt and refined the model with good stereochemistry using the programs Coot¹⁷ and Phenix¹⁵ (Supplementary Table 1). All figures of structural models were generated using Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

Determining affinities of WT and mutant nsp16/nsp10 binding to magnesium

We used microscale thermophoresis (MST) to derive equilibrium dissociation constants (K_D) for protein-metal interactions. A detailed protocol has been published⁸. Briefly, $20\text{ }\mu\text{M}$ of each enzyme complex was labeled by incubating with dye solution ($60\text{ }\mu\text{M}$) in the labeling buffer at room temperature for 30 min. The magnesium stock was prepared in the MST reaction buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5% glycerol, and 0.05% Tween 20), and two-fold serial dilutions (from 2 mM stock) were made in 12 steps. The labeled protein (20 nM) was equally mixed into each ligand reaction (ligand concentration ranges 500 nM to 1 mM). The final reaction mixtures were loaded and measured on a Monolith NT.115 instrument (NanoTemper Technologies) at 25°C . The results shown here are from three independent experiments. Data were fitted by a single-site binding model in GraphPad Prism (GraphPad Software, San Diego, CA).

Enzyme activity assay

We used a LC/MS-based method (reported earlier)⁸ for quantitative measurement of Cap-1 formation by nsp16/nsp10 enzymes in the presence of various metals. Briefly, 0.1 μM of enzymes were allowed to react with 1 μM ^{me7}GpppA-capped 25 nt RNA in a buffer (50 mM Tris-HCl, pH 8.0, 5 mM KCl, 1 mM DTT, 0.2 mM SAM) supplemented with 1 mM MgCl₂ or MnCl₂ or CaCl₂ or 5 mM EDTA. The reactions were incubated at 37°C for 30 min, stopped by heating at 75°C for 5 min in the presence of 5 mM EDTA, and were subjected to LC/MS intact mass analysis. Nucleic acids in the samples were separated using a Thermo DNAPac™ RP Column on a Vanquish Horizon UHPLC System, followed by mass determination using a Thermo Q-Exactive Plus mass spectrometer. The raw data were deconvoluted using Promass HR (Novatia, LLC). The deconvoluted mass peak ratios between reactants and the expected products was used to estimate the percentage of 2'-O methylation. Results shown in Fig. 2e are average of three independent experiments (n=3) normalized to a WT dataset consisted of seven data points. Source data are provided as a Source Data File. The RNA substrate used in this assay is ^{me7}GpppAUAGAACUUCGUCGAGUACGCUCAA-[6-FAM].

DATA AVAILABILITY

The information about coding sequences of nsp16 (NCBI reference sequence YP_009725311.1) and nsp10 (NCBI reference sequence: YP_0009725306.1) of the seafood market pneumonia SARS-CoV-2 isolate Wuhan-Hu-1 (NC_045512) used in this study is available at NCBI (https://www.ncbi.nlm.nih.gov/nuccore/NC_045512). Files for atomic coordinates and structure factors were deposited in the Protein Data Bank under accession codes 7LW3 (product bound) and

7LW4 (SAH bound). Correspondence and requests for material should be addressed to Y.K.G. (guptay@uthscsa.edu).

REFERENCES

1. Decroly, E., Ferron, F., Lescar, J. & Canard, B. Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* **10**, 51-65 (2011).
2. Bouvet, M. et al. In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog* **6**, e1000863 (2010).
3. Chen, Y. et al. Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog* **7**, e1002294 (2011).
4. Daffis, S. et al. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* **468**, 452-6 (2010).
5. Habjan, M. et al. Sequestration by IFIT1 impairs translation of 2'O-unmethylated capped RNA. *PLoS Pathog* **9**, e1003663 (2013).
6. Züst, R. et al. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol* **12**, 137-43 (2011).
7. Zhou, Z. et al. Heightened Innate Immune Responses in the Respiratory Tract of COVID-19 Patients. *Cell Host Microbe* **27**, 883-890 e2 (2020).
8. Viswanathan, T. et al. Structural basis of RNA cap modification by SARS-CoV-2. *Nat Commun* **11**, 3718 (2020).
9. Rosas-Lemus, M. et al. High-resolution structures of the SARS-CoV-2 2'-O-methyltransferase reveal strategies for structure-based inhibitor design. *Sci Signal* **13**(2020).
10. Decroly, E. et al. Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS Pathog* **7**, e1002059 (2011).
11. Butler, D.J. et al. Shotgun Transcriptome and Isothermal Profiling of SARS-CoV-2 Infection Reveals Unique Host Responses, Viral Diversification, and Drug Interactions. *bioRxiv* (2020).
12. Di Filippo, L. et al. Hypocalcemia is highly prevalent and predicts hospitalization in patients with COVID-19. *Endocrine* **68**, 475-478 (2020).
13. Cheungpasitporn, W. et al. Impact of admission serum calcium levels on mortality in hospitalized patients. *Endocr Res* **43**, 116-123 (2018).
14. Zeng, H.L., Yang, Q., Yuan, P., Wang, X. & Cheng, L. Associations of essential and toxic metals/metalloids in whole blood with both disease severity and mortality in patients with COVID-19. *FASEB J* **35**, e21392 (2021).
15. Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-3 (1994).
16. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
17. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).

Acknowledgments

This work was partially supported by funding from the San Antonio Partnership for Precision Therapeutics (SA-PPT), pilot awards from the Institute for Integration of Medicine and Science (IIMS) and UT Health San Antonio (UTHSA), and the Max and Minnie Tomerlin Voelcker Foundation to Y.K.G. T.V. is supported by a research training award (RP170345) from the Cancer Prevention Research Institute of Texas (CPRIT). Y.K.G. is also supported by a high impact/high risk award from the CPRIT, and a Rising STARS award from the UT System. We are grateful to beamline scientists at NECAT-24ID sector, APS, Chicago for provision of synchrotron beamtime and data collection. This work is based on research conducted at the Northeastern Collaborative Access Team beamlines (NIH P30GM124165, and DOE DE-AC02-06CH11357 grants). We also thank the X-ray Core of UTHSA.

Author contributions

Y.K.G. conceived, designed, and supervised the overall study, and performed crystallographic studies; T.V., A.M., S.A. purified proteins, T.V. performed crystallization, S.Q. and T.V. performed biochemical assays. S.-H.C. performed the LC/MS-based assays with assistance from N.D. L.M.-S. provided reagents. Y.K.G. wrote the manuscript, and all authors have read and approved this version.

Competing interests

Y.K.G is founder of Atomic Therapeutics. S-H. C. and N.D. are employees of New England Biolabs, a manufacturer and vendor of molecular biology reagents, including vaccinia RNA

capping enzyme and cap 2'-*O* methyltransferase. None of these affiliations affect the authors' impartiality, adherence to journal standards and policies, or availability of data.

Materials and correspondence

Files for atomic coordinates and structure factors were deposited in the Protein Data Bank under accession codes 7LW3 (product bound) and 7LW4 (SAH bound). Correspondence and requests for material should be addressed to Y.K.G. (guptay@uthscsa.edu).

FIGURE LEGENDS

Fig. 1. Structures of SARS-CoV-2 nsp16/nsp10 complexes. **a**, The substrate ($^{m^7}GpppA$, blue stick) and methyl donor (SAM, yellow stick)-bound nsp16 (cyan)/nsp10 (orange) complex (PDB ID, 6WKS)⁸ represents a closed form. **b**, The product ($^{m^7}GpppAmU$, red stick; byproduct SAH [grey stick])-bound nsp16 (blue)/nsp10 (magenta) in an *open* state. A yellow circle shows the methylated ribose (2'-O-me) of N1 (A) base. **c**, The SAH (grey) bound nsp16 (grey)/nsp10 (pink) represents a *partially open* or *enzyme reset* state. **d**, Secondary structure-based overlay of nsp16 in substrate- and product-bound states clearly shows the universal expansion of the enzyme upon 2'-O methylation. **e**, A close up view of Cap-binding and catalytic pocket of the product structure shows nsp16 residues (cyan sticks) interacting with Cap-1 (red). A positional change in orientation of the substrate (Cap-0, blue) from the “closed” structure determined previously⁸ is shown. **f**, An overlay of the product (Cap-1) and byproduct (SAH)-bound structures shows change in the orientation of gate loop 2. Reduction in buried surface area between nsp16/nsp10 in *fully* and *partially open* structures (compared to substrate-bound closed state) is shown (**g-i**).

Fig. 2. Metal dependency of nsp16/nsp10 and its clinical variant for 2'-O methylation.

a, Alignment of nsp16 from different CoV representing all three sub-classes (α , β , γ). Blue sphere denotes S33R, which locates in gate loop 1; S33 is an asparagine (N) in SARS-CoV; S33R is a clinical nsp16 variant of SARS-CoV-2; Cyan triangle, N198 that coordinates Mg^{2+} ; black sphere, catalytic lysine (K46 of KDKE tetrad). **b**, A close-up view of coordination of Mg^{2+} ion (green sphere). Mg^{2+} coordinates with five water molecules (black spheres) at the nucleic acid face and the side chain of N198 on the opposite face. Two of these water molecules hold the phosphoryl oxygens of the U_2 . Red arrow; 2'-O-methylated ribosyl of A_1 base. **c**, The side chain of arginine

(magenta, modeled) at the S33 (cyan) position intrudes into the Mg^{2+} pocket and thus, may displace Mg^{2+} or disrupt the Mg^{2+} /water network. **d**, Binding isotherms of Mg^{2+} nsp16/nsp10 (WT and mutants) interaction derived from the MST data. (N.D., not determined). **e**, Quantitative measurement of Cap-1 formation by nsp16/nsp10 enzymes experiments (+/- metals, EDTA) as derived from the LC/MS data. Results are averages of three independent experiments (n=3) normalized to a WT dataset consisted of seven datapoints with one standard deviation (s.d.) for each metal or EDTA shown as error bars. Source data are provided as a Source Data File. **no_enz**, reaction devoid of nsp16/nsp10 enzyme. **f**, An overlay of the N_1 and N_2 bases and SAH of SARS-CoV-2 nsp16/nsp10 (red) and dengue NS5 (blue; PDB ID: 5DTO) RNA caps shows entirely different orientations of the terminal base of the cap ($^{m^7}G$), two phosphates (β and γ) and Mg^{2+} ions. Mg^{2+} (in dengue, yellow sphere) stabilizes the three phosphates whereas in SARS-CoV-2 (green sphere) it indirectly (water-mediated) stabilizes the phosphate of the N_2 base on one side and engages N198 of nsp16 on the opposite side. Red arrow; 2'-*O* methyl of ribose of N_1 .

Figure 1

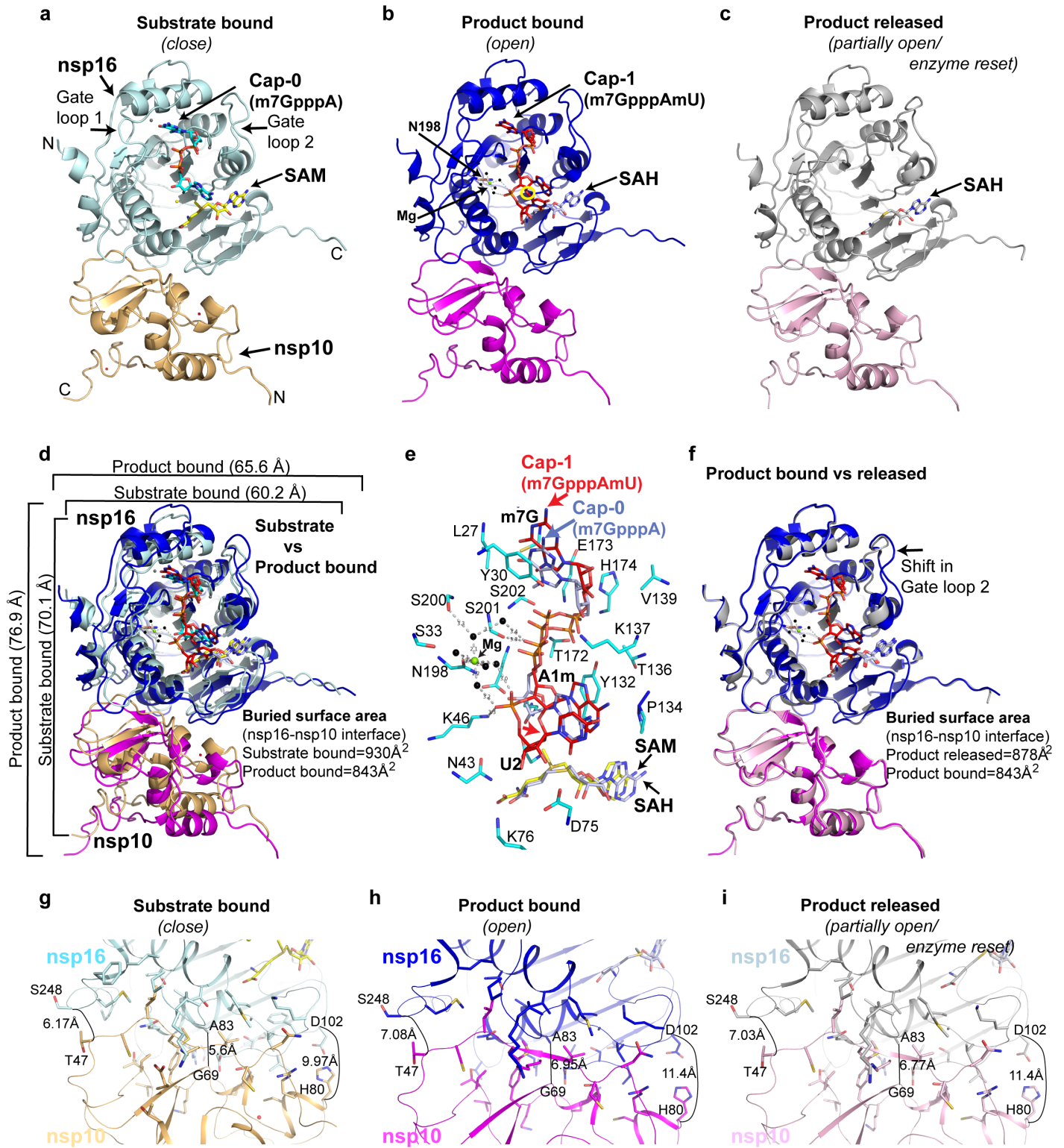
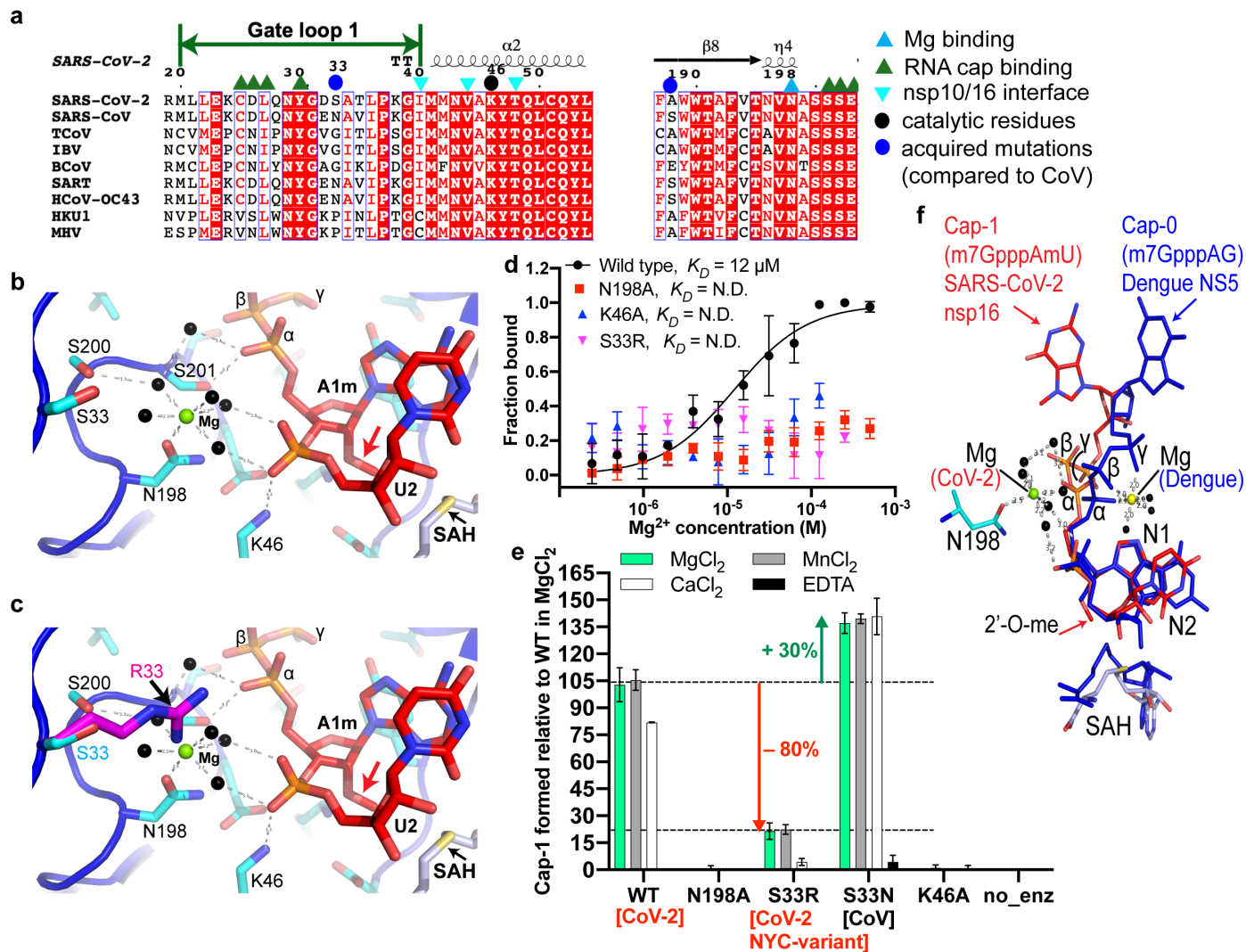


Figure 2



SUPPLEMENTARY INFORMATION

A metal ion orients the mRNA to ensure accurate 2'-O ribosyl methylation of the first nucleotide of the SARS-CoV-2 genome

Thiruselvam Viswanathan^{1,2,#}, Anurag Misra^{1,2,#}, Siu-Hong Chan³, Shan Qi^{1,2}, Nan Dai³, Shailee Arya¹, Luis Martinez-Sobrido⁴, Yogesh K. Gupta^{1,2,6,*}

¹ Greehey Children's Cancer Research Institute

² Department of Biochemistry and Structural Biology

University of Texas Health at San Antonio

8403 Floyd Curl Drive, San Antonio, TX 78229, USA

³ New England Biolabs, 240 County Road, Ipswich, MA, 01938, USA

⁴ Texas Biomedical Research Institute, San Antonio, TX, 78227, USA

⁶ Lead contact

*Corresponding author:

Y.K.G email: guptay@uthscsa.edu

These authors contributed equally

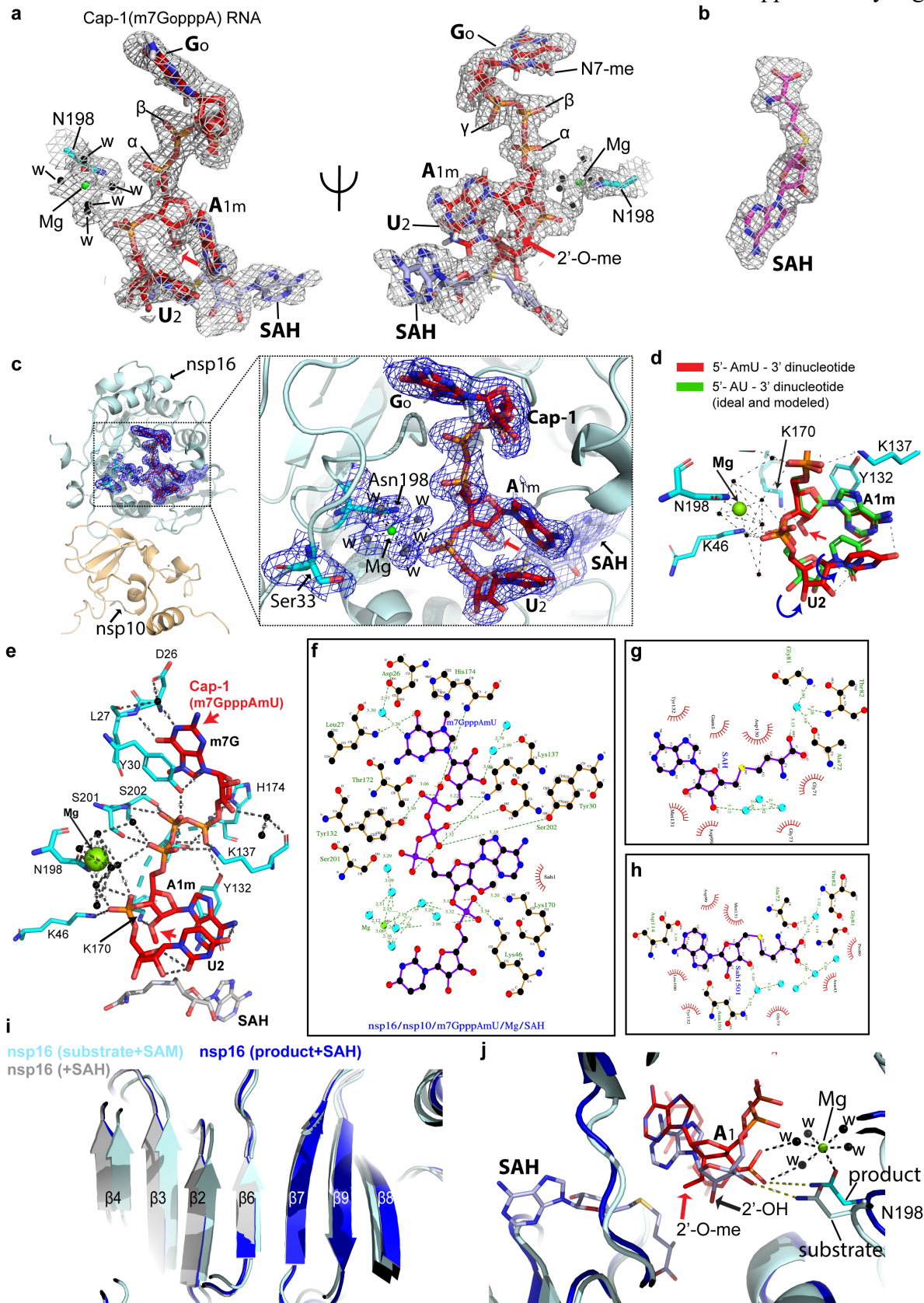
This document contains the following:

Supplementary Figure 1

Supplementary Table 1

Supplementary References

Supplementary Figure 1



S. Fig. 1. a, A close-up view of the product (m^7 GpppAmU, red stick), SAH (grey stick), magnesium (green sphere), water (W, black spheres), and N198 binding. The Fo-Fc electron density omit map for these ligands contoured at 2.7σ is shown as a grey mesh. To minimize the possibility of bias introduced by ligands, all ligands were excluded from refinement and phase calculations. Positions of the methyl group attached to the acceptor moiety (2'-O in ribose of the target nucleotide A_{1m}) is depicted by red arrow. **b,** The Fo-Fc electron density omit map for SAH (red stick) in nsp16/nsp10/SAH structure contoured at 2.7σ is shown as a grey mesh. **c,** Final structure of nsp16 (cyan)/nsp10(orange)/ m^7 G_opppAmU (red)/SAH (grey)/ Mg^{2+} (green)/ Mg^{2+} -coordinating waters showing 2Fo-Fc map (blue mesh) contoured at 1.0σ . **d,** An overlay of an A in an ideal AU (green) dinucleotide over Am base in the product structure shows deviated geometry (outward motion, blue arrows) of the U2 base, suggestive of a state preceding the product's release. **e,** nsp16 residues (cyan) that directly interact with Cap-1 RNA (red) and Mg^{2+} (green sphere) and water (black sphere). **f,** Protein-ligand interaction network of nsp16/nsp10/Cap-1 (m^7 GpppAmU) and SAH in product (**g**), and SAH-bound (**h**) nsp16/nsp10 structures. The green dashed lines represent hydrogen bonding, and the cyan spheres represent water molecules. These figures were generated using the LigPlot+ program¹. **i,** A secondary structure-based superposition of nsp16 in all three structures show good overlay of the central β -sheet in all three structures. The regions flanking the central core in nsp16 and the entire nsp10 universally expands (relative to substrate/SAM-bound form, cyan) to assume a more relaxed (or fully open state) in the product-bound form (blue). The structures of product plus SAH and only SAH-bound enzymes do not deviate except in the gate loop region. **j,** An overlay of the substrate (light cyan) and product (blue) structure is shown with their respective Caps (Cap1U as red stick in product, grey stick; Cap-0 in substrate structures). In the absence of Mg^{2+} in the substrate structure, the side chain of N198 coordinates with 3'-OH of the N1 base whereas in the product structure it coordinates with the phosphoryl oxygen of the N2 base and Mg^{2+} ion.

Supplementary Table 1.

Data collection and refinement statistics (molecular replacement)

	nsp16/nsp10/m7GpppAmU (PDB ID: 7LW3)	nsp16/nsp10/SAH (PDB ID: 7LW4)
Data Collection	NECAT-24ID, APS	NECAT-24ID, APS
Wavelength	1.071	1.071
Resolution range (Å)	28.63 - 2.3	29.75 - 2.5
Space group	P3 ₁ 21	P3 ₁ 21
Unit cell (Å)	a=b, c (Å)	
	184, 57.1	184.5, 56.9
	$\alpha=\beta, \gamma$ (°)	
	90, 120	90, 120
Total reflections	98827 (9874)	76521 (7646)
Unique reflections	49419 (4937)	38625 (3842)
Multiplicity	21.13 (3.75)	2.0 (2.0)
Completeness (%)	99.9 (100)	99.79 (100)
Mean I/sigma(I)	15.2 (0.9)	11.80 (3.02)
Wilson B-factor	44.93	47.39
R-merge	0.089 (2.34)	0.113 (1.41)
Refinement		
Reflections used in refinement	49411 (4937)	38620 (3842)
R-work	0.22 (0.28)	0.20 (0.26)
R-free	0.25 (0.32)	0.24 (0.29)
Number of non-hydrogen atoms	3444	3384
	macromolecules	3192
	ligands	91
	solvent	161
RMS (bonds)	0.027	0.034
RMS (angles)	2.37	2.43
Ramachandran favored (%)	93.64	94.61
Ramachandran allowed (%)	5.62	4.9
Ramachandran outliers (%)	0.73	0.49
Average B-factor (Å ²)	57.52	58.58
	macromolecules	57.31
	ligands	67.12
	solvent	56.3

*Values for outermost shell are given in parentheses.

Supplementary References.

1. Laskowski R. A., Swindells M. B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.*, **51**, 2778-2786.