Recognition and Cleavage of Human tRNA Methyltransferase TRMT1 by the SARS-CoV-2 Main Protease

Angel D'Oliviera[1], Xuhang Dai[2], Saba Mottaghinia[3], Evan P. Geissler[1], Lucie Etienne[3], Yingkai Zhang[2,4], Jeffrey S. Mugridge[1]*

[1] Department of Chemistry & Biochemistry, University of Delaware, Newark, DE 19716.
[2] Department of Chemistry, New York University, New York, NY 10003.
[3] CIRI (Centre International de Recherche en Infectiologie), Univ Lyon, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, ENS de Lyon, F-69007 Lyon, France
[4] Simons Center for Computational Physical Chemistry at New York University, New York, NY 10003.

* Correspondence and requests for materials should be addressed to J.S.M. (email: mugridge@udel.edu)

Author Contributions: J.S.M designed the research. E.P.G. provided preliminary data for research design. A.D. and J.S.M. performed cloning, protein expression, purification, crystallography and structure determination, peptide cleavage kinetic assays, mammalian cell culture, and proteolysis assays. X.D. and Y.Z. performed molecular dynamics simulations. S.M. and L.E. performed evolutionary analysis. J.S.M. and A.D. wrote the manuscript and J.S.M., A.D., X.D., S.M., L.E., and Y.Z. edited the manuscript.

Competing Interest Statement: The authors declare no competing interest.

Abstract

The SARS-CoV-2 main protease (M^{pro}) is critical for the production of functional viral proteins during infection and, like many viral proteases, can also target host proteins to subvert their cellular functions. Here, we show that the human tRNA methyltransferase TRMT1 can be recognized and cleaved by SARS-CoV-2 M^{pro}. TRMT1 installs the N2,N2-dimethylguanosine (m2,2G) modification on mammalian tRNAs, which promotes global protein synthesis and cellular redox homeostasis. We find that M^{pro} can cleave endogenous TRMT1 in human cell lysate, resulting in removal of the TRMT1 zinc finger domain required for tRNA modification activity in cells. Evolutionary analysis shows that the TRMT1 cleavage site is highly conserved in mammals, except in Muroidea, where TRMT1 may be resistant to cleavage. In primates, regions outside the cleavage site with rapid evolution could indicate adaptation to ancient viral pathogens. We determined the structure of a TRMT1 peptide in complex with M^{pro}, revealing a substrate binding conformation distinct from the majority of available M^{pro}-peptide complexes. Kinetic parameters for peptide cleavage showed that the TRMT1(526-536) sequence is cleaved with comparable efficiency to the M^{pro}-targeted nsp8/9 viral cleavage site. Mutagenesis studies and molecular dynamics simulations together indicate that kinetic discrimination occurs during a later step of M^{pro}-mediated proteolysis that follows substrate binding. Our results provide new information about the structural basis for M^{pro} substrate recognition and cleavage that could help inform future therapeutic design and raise the possibility that proteolysis of human TRMT1 during SARS-CoV-2 infection suppresses protein translation and oxidative stress response to impact viral pathogenesis.

Significance Statement

Viral proteases can strategically target human proteins to manipulate host biochemistry during infection. Here, we show that the SARS-CoV-2 main protease (M^{pro}) can specifically recognize and cleave the human tRNA methyltransferase enzyme TRMT1, which installs a modification on human tRNAs that is critical for protein translation. Our structural and functional analysis of the M^{pro}-TRMT1 interaction shows how the flexible M^{pro} active site engages a conserved sequence in TRMT1 in an uncommon binding mode to catalyze its cleavage and inactivation. These studies provide new insights into substrate recognition by SARS-CoV-2 M^{pro} that could inform future antiviral therapeutic design and suggest that proteolysis of TRMT1 during SARS-CoV-2 infection may disrupt tRNA modification and host translation to impact COVID-19 pathogenesis or phenotypes.

Introduction

SARS-CoV-2 has resulted in over 6 million deaths worldwide since the start of the COVID-19 pandemic in early 2020 (1, 2). The development of mRNA and other vaccines has played a large and critical role in reducing mortality since their introduction in 2021 (3–6), but a fundamental understanding of coronavirus host-protein interactions and biology continues to be an important goal to inform ongoing and future therapeutic design. The SARS-CoV-2 main protease (M^{pro}) is a well-studied antiviral drug target because its activity is essential for viral replication. More than 600 M^{pro} crystal structures – the majority with bound small molecule inhibitors – are currently available in the protein databank (7). M^{pro} is necessary for the proteolysis of 11 different cleavage sites in the two SARS-CoV-2 extended viral polypeptides, which results in the liberation of mature non-structural proteins (nsps) that are essential for host infection and viral propagation (8–13). Approximately twenty structures of M^{pro} in complex with different viral cleavage sequences have been determined (14–16), in which the M^{pro} homodimer has a peptide substrate bound to its active site (**Figure 1A**). However, a detailed understanding of SARS-CoV-2 M^{pro} substrate selectivity and cleavage efficiency remains poorly defined, and the structural characterization of interactions between the viral protease and human host protein targets remains underexplored (17).

Early in the COVID-19 pandemic, Gordon *et al.* expressed tagged SARS-CoV-2 proteins in human embryonic kidney cells (HEK293T/17) and employed affinity purification mass spectrometry (AP-MS) to map proteome-wide virus-host protein interactions, in which a putative interaction between catalytically inactive SARS-CoV-2 M^{pro} (Cys145Ala) and the human tRNA methyltransferase TRMT1 was identified (18). Analogous AP-MS experiments with wild-type M^{pro} found no stable interaction with TRMT1, and subsequent mapping with SARS-CoV-1 proteins also found a TRMT1 interaction exclusively with Cys145Ala M^{pro}, suggesting that coronavirus M^{pro} may recognize and actively cleave human TRMT1 in cells (18, 19). TRMT1 is a tRNA-modifying enzyme responsible for installing N^2,N^2-dimethylguanosine (m2,2G), an abundant tRNA modification found at the G26 position of human tRNAs (20). Human HEK293T cells lacking TRMT1, and therefore the m2,2G26 modification, have significantly decreased global protein synthesis and reduced proliferation (21). Human neural stem cells with TRMT1 knockdowns were found to have hypersensitivity to redox stress, implicating TRMT1 and the m2,2G26 modification in the regulation of redox homeostasis (21). SARS-CoV-2 M^{pro}-directed cleavage of TRMT1, and subsequent downregulation of tRNA m2,2G26 modification during infection, could therefore have direct impacts on both host and viral protein synthesis, as well as phenotypes linked to redox stress.

Here, we show that SARS-CoV-2 M^{pro} proteolyzes endogenous human TRMT1 in human cell lysate and cleaves an amino acid sequence located between methyltransferase and zinc finger domains in TRMT1 with similar kinetic parameters to known viral polypeptide cleavage sites. We have determined the structure of M^{pro} in complex with a TRMT1 peptide that shows how the viral protease recognizes and cleaves this human protein sequence and reveals a unique binding mode for the TRMT1 peptide in the M^{pro} active site. Further kinetic analysis and molecular dynamics simulations show how the flexible M^{pro} active site can accommodate and cleave diverse peptide sequences and geometries, and suggest that kinetic discrimination between substrate sequences occurs during later steps in M^{pro}-catalyzed amide bond cleavage reaction. Like many viruses, SARS-CoV-2 hijacks and disrupts diverse host biochemical pathways thereby promoting viral replication and avoiding host detection and immunological response. Recent studies have identified a large number of human proteins that may be targeted and cleaved by M^{pro}, many of which are involved in regulation of cellular proliferation, inflammatory response, transcription, translation, ubiquitination, apoptosis, and metabolism (22–28). Our biochemical and structural data show that TRMT1 can be targeted and cleaved by the SARS-CoV-2 main protease. SARS-CoVs employ multiple strategies to alter host translation (29, 30), and cleavage of TRMT1 and corresponding loss of the tRNA m2,2G26 modification may be another mechanism by which the virus is able to disrupt regulation of host protein synthesis. Furthermore, it is possible that proteolysis of TRMT1

during SARS-CoV-2 infection could contribute to some cellular pathogenesis observed with COVID-19.

Results

Full-length TRMT1 is cleaved by SARS-CoV-2 M^{pro} *in vitro*. SARS-CoV-1 and SARS-CoV-2 proteome-wide virus-host protein interaction maps identified a putative interaction between catalytically inactive M^{pro} (Cys145Ala) and human TRMT1, but no such stable interaction was found using wild-type protease, suggesting M^{pro} may target and cleave TRMT1 in cells (18, 19). Human TRMT1 contains a peptide sequence (527 – 534) consistent with the cleavage consensus for M^{pro}, located in an AlphaFold-predicted (31, 32) surface-exposed region (**Figure S1**) linking the TRMT1 N-terminal SAM-dependent methyltransferase (MTase) domain and the C-terminal zinc finger (ZF) domain (**Figure 1B**). To determine if full-length (FL) TRMT1 is susceptible to cleavage by M^{pro}, we first measured proteolysis of human TRMT1 expressed and enriched from *E. coli* by recombinant purified SARS-CoV-2 M^{pro} (**Figure 1C**). TRMT1 was incubated with either catalytically inactive M^{pro} Cys145Ala or active wild-type M^{pro}, and TRMT1 cleavage was monitored by Western-blot using two TRMT1-specific polyclonal antibodies: a dual-domain recognizing antibody specific for portions of both the methyltransferase and zinc finger domain (anti-TRMT1 460-659), and a single-domain recognizing antibody specific for only the zinc finger domain (anti-TRMT1 609-659). As expected, FL TRMT1 (~75kDa) was stable during a 2-hour incubation with inactive Cys145Ala M^{pro}. However, when incubated with wild-type M^{pro}, the intensity of the band corresponding to FL TRMT1, as measured by both TRMT1 antibodies, was reduced by >90% after 2 hours. In Western blots with cleaved TRMT1, the dual domain-recognizing anti-TRMT1(460-659) antibody showed the appearance of two new lower molecular weight bands, corresponding to the anticipated molecular weights of the M^{pro}-mediated TRMT1 cleavage products (~61 kDa TRMT1 methyltransferase domain and ~14 kDa TRMT1 zinc finger domain). After demonstrating that recombinant FL TRMT1 isolated from *E. coli* could be cleaved by SARS-CoV-2 M^{pro}, we next asked whether endogenous TRMT1 from human cells could be similarly cleaved by the viral protease. HEK293T cell lysate was incubated with either Cys145Ala or WT M^{pro}, and endogenous TRMT1 levels were monitored over time by Western blot (**Figure 1D**). Incubation with Cys145Ala M^{pro} resulted in no change to the FL TRMT1 band over the course of 10 hours, while incubation with wild-type M^{pro} resulted in time-dependent proteolysis of FL TRMT1. Unlike in cleavage assays with recombinant TRMT1, no build-up of cleavage products was observed during M^{pro}-mediated proteolysis of endogenous TRMT1, suggesting instability and rapid degradation of cleaved TRMT1 fragments in human cell lysate. Together, these experiments demonstrate that both recombinant FL TRMT1 and endogenous FL TRMT1 in human cell lysate are viable substrates for cleavage by SARS-CoV-2 M^{pro}.

Structure of TRMT1 peptide bound to the M^{pro} active site. To visualize how TRMT1 is recognized and cleaved by the viral protease active site, we determined the co-crystal structure of catalytically inactive Cys145Ala mutant M^{pro} in complex with the TRMT1 peptide sequence corresponding to the expected cleavage site, human TRMT1 residues 526-536 (**Figure 2**, **Table S1**). The 1.9 Å resolution structure shows clear density for the TRMT1 peptide bound to the active site in one protomer of the M^{pro} dimer (**Figure 2A**, **Figure S2**). As expected based on alignment with the M^{pro} consensus cleavage sequence, TRMT1 Gln530 occupies the critical S1 subsite pocket in the M^{pro} active site (**Figure 2A**), positioning the Gln530-Ala531 peptide bond directly adjacent to the His41/Cys145 catalytic dyad for proteolytic cleavage (**Figure 2B**). TRMT1 Leu529 corresponds to substrate position P2 and occupies the hydrophobic M^{pro} S2 pocket. TRMT1 peptide binding to M^{pro} is mediated by key hydrogen bonding contacts, including M^{pro} residues Phe140, His163, and Glu166 that specifically recognize TRMT1 Gln530, in addition to other interactions including M^{pro} Thr24-TRMT1 Thr534, M^{pro} Thr26-TRMT1 Asn532, M^{pro} Asn142-TRMT1 Asn532, M^{pro} Glu166-TRMT1 Arg528, and M^{pro} Gln189-TRMT1 Leu529, which position the TRMT1 peptide in the M^{pro} active site. These primary hydrogen bonding contacts are consistent with previous structures of M^{pro} bound to viral peptide substrates (14). Importantly, the identified M^{pro}-targeted residues in

human TRMT1 are conserved in the human population (*i.e.*no missense polymorphisms), showing that human TRMT1 can be recognized and cleaved by SARS-CoV-2 M^{pro}.

TRMT1 engages the M^{pro} active site in a distinct binding conformation. M^{pro} has more than 10 native viral polypeptide substrates with relatively high sequence variability outside the conserved P1 Gln and the adjacent P2 and P1′ residues. We compared our human TRMT1-bound M^{pro} structure with available SARS-CoV-2 M^{pro} structures bound to known viral polypeptide cleavage sequences (**Figure 3A**; nsp4/5, nsp5/6, nsp6/7, nsp8/9, nsp9/10, nsp10/11, and nsp15/16). The overall structures of M^{pro} are very similar for all peptide-bound structures, with M^{pro} backbone RMSD values all below 1.6 (**Figure S3**). Analysis of aligned peptide substrates in the M^{pro} active site shows that both TRMT1 and viral peptide backbones have nearly identical conformations for the N-terminal P4 → P1′ residues. In contrast, backbone geometries of C-terminal peptide residues (P2′ → P4′) diverge more significantly and fall into two distinct binding modes distinguished primarily by the P2′ Ψ dihedral angle (**Figure 3B, C**). The majority of viral peptide substrates adopt a binding conformation with P2′ $\Psi \geq 157°$ in which the P3′ sidechain is positioned away from the M^{pro} surface (**Figure 3D**), which we have designated as the 'P3′-out' conformation. In contrast, TRMT1 and a single viral peptide substrate, nsp6/7, bind in a distinct conformation with P2′ $\Psi \leq 116°$ (**Figure 3C**) in which the P3′ sidechain (Phe in TRMT1, Met in nsp6/7) is positioned toward the M^{pro} surface where it displaces M^{pro} Met49 to open and occupy pocket S3′ (**Figure 3D**). We have designated this uncommon binding mode as the 'P3′-in' conformation.

M^{pro} cleaves TRMT1 peptide with comparable kinetics to known viral polypeptide cleavage sites. To further examine whether TRMT1 is a viable substrate for SARS-CoV-2 M^{pro}, we compared kinetic parameters for proteolysis of the TRMT1 (526-536) peptide and the native N-terminal auto-processing sequence nsp4/5, using fluorogenic peptide cleavage assays. The TRMT1 peptide is cleaved noticeably slower than the nsp4/5 auto-processing sequence (**Figure 4A, Table S2**), with an approximately 200-fold decrease in k_{cat}, 4-fold decrease in K_M, and 50-fold decrease in catalytic efficiency compared to M^{pro}-mediated cleavage of the nsp4/5 substrate (**Figure 4B, Table S2**). However, our measured catalytic efficiency for M^{pro}-mediated TRMT1 cleavage (1.82×10^{-4} $\mu M^{-1}s^{-1}$) is within two-fold of that recently reported by MacDonald *et al.* for cleavage of the known viral nsp8/9 cleavage sequence (3.6×10^{-4} $\mu M^{-1}s^{-1}$).(15) Although the cleavage assay conditions used by MacDonald *et al*. differ slightly from our assay conditions, we measure nearly identical cleavage kinetics for nsp4/5 (9.7×10^{-3} $\mu M^{-1}s^{-1}$ vs. 1.26×10^{-2} $\mu M^{-1}s^{-1}$), suggesting TRMT1 and nsp8/9 kinetic parameters are comparable. These kinetic experiments show that, while the TRMT1 M^{pro} target sequence is cleaved less efficiently than the canonical nsp4/5 auto-processing site, the TRMT1 sequence is cleaved with kinetics very similar to the known nsp8/9 M^{pro} cleavage site in the viral polypeptide. This suggests that TRMT1(526-536) is a viable substrate for M^{pro}-targeted proteolysis.

M^{pro} residues involved in TRMT1 substrate recognition have only small effects on cleavage efficiency. We next compared the M^{pro}-peptide interactions observed in TRMT1-, nsp4/5-, and nsp8/9-bound structures and identified several M^{pro} residues involved in direct substrate contacts that differ between the TRMT1- and nsp4/5- or nsp8/9-bound structures (**Figure 4C**). In the TRMT1-bound structure, Met49 is shifted to accommodate P3′ substrate residue Phe533 binding in the S3′ pocket, and the M^{pro} Asn142 and Gln189 sidechains are repositioned to form hydrogen bonds with the TRMT1 peptide backbone. To test how these M^{pro} active site residues affect TRMT1 recognition and catalysis, we generated single alanine point mutations at each of these three sites and measured kinetic parameters for M^{pro}-mediated proteolysis of nsp4/5 and TRMT1 peptides. Surprisingly, we saw no substantial changes in proteolysis kinetics for either nsp4/5 or TRMT1 substrates with any of the tested M^{pro} mutants (**Figure 4D, Table S2**), suggesting none of these residues play a significant role in determining differential cleavage rates for TRMT1 vs nsp4/5.

TRMT1 mutations predicted to alter peptide binding conformation have little effect on cleavage kinetics. Our TRMT1-bound M^{pro} structure shows that in the unique P3′-in conformation, the TRMT1 P1′ Ala side chain is positioned within ~3 Å of the P3′ Phe ring in the S3′ pocket. A

bulkier residue in the P1′ position of TRMT1 would clash with this P3′ Phe orientation and disfavor the P3′-in binding mode. To test whether mutations that likely disfavor the P3′-in conformation would affect TRMT1 recognition and proteolysis, we mutated TRMT1 P1′ Ala to the viral P1′ consensus of Ser, and measured kinetic parameters for cleavage of the WT and Ala531Ser TRMT1 peptides. In these experiments, we saw no major change in cleavage kinetics for the Ala to Ser mutation at the P1′ position in TRMT1 (**Table S2, Figure S4**), suggesting that disfavoring TRMT1's P3′-in binding conformation does not substantially alter M^{pro}'s ability to cleave the peptide substrate. Together, our mutagenesis and kinetic experiments show that mutations to apparently key residues in either the peptide substrate or the M^{pro} active site have little effect on proteolysis activity. Consistent with other recent kinetic studies (15, 33), these data suggest M^{pro} has a flexible active site that can accommodate cleavage of diverse substrate sequences and substrate geometries.

Molecular dynamics simulations suggest kinetic discrimination happens during later steps of M^{pro}-catalyzed substrate cleavage. To further support the structural and biochemical data described above and better understand how peptide substrate binding geometry may or may not be linked to M^{pro}-catalyzed cleavage efficiency, we carried out molecular dynamics (MD) simulations using nsp4/5-, nsp8/9-, and TRMT1-bound M^{pro} complexes. MD simulations show TRMT1 is stably-bound to the M^{pro} active site and primarily adopts the P3′-in conformation observed in the crystal structure, where TRMT1 Phe533 occupies the S3′ pocket with a probability of ~65% over the course of the 100 ns simulation (**Figure 5A**). In contrast, simulations of the nsp4/5-M^{pro} complex show the nsp4/5 peptide primarily adopts the P3′-out conformation, where nsp4/5 Phe531 is oriented away from the M^{pro} surface with ~70% probability, consistent with the observed nsp4/5 binding mode in the crystal structure. This computational analysis shows that while P3′-in and -out conformations can interconvert, the dominant binding modes for TRMT1 and nsp4/5 peptide substrates (P3′-in vs -out, respectively) are consistent across the crystal structures and MD simulations. Furthermore, because these conformations differ primarily at the C-terminal end of the peptide, but have very similar geometries at the scissile amide bond, we find that the computationally and experimentally determined substrate binding poses fail to explain the large differences in cleavage kinetics observed for TRMT1 versus nsp4/5.

Previous structural and geometric analysis of diverse serine protease substrate-inhibitor complexes has shown than the consensus nucleophilic attack angle in the reactive Michaelis complex is approximately 90° (34). We next asked whether differences in nucleophilic attack angle measured over the course of the MD simulation might be able to explain changes in M^{pro}-mediated cleavage kinetics. The distribution of the S–C=O nucleophilic attack angle shows that the M^{pro}-nsp4/5 complex has only ~9% increased probability to fall within 89± 7°, as compared with the M^{pro}-TRMT1 complex over the course of the simulation (**Figure 5B**). Thus, while the nsp4/5 peptide shows slightly more favorable positioning for nucleophilic attack in the Michaelis complex during the first step of cleavage, this is not nearly sufficient to explain the large difference (~200-fold in k_{cat}) in experimentally observed cleavage kinetics between nsp4/5 and TRMT1. Taken together with our TRMT1-M^{pro} structure and kinetic analysis of M^{pro} and TRMT1 mutants, these results suggest that kinetic discrimination between peptide cleavage rates is likely to occur during a later step in the M^{pro}-catalyzed cleavage reaction that follows substrate binding, Michaelis complex formation, and initial nucleophilic attack, similar to some serine proteases (34).

Discussion

Recent studies mapping host-coronavirus protein interaction networks predicted high confidence interactions between the SARS-CoV-1 or SARS-CoV-2 main protease (M^{pro}) and the human tRNA-modifying enzyme TRMT1 (18, 35). In this work, we show that SARS-CoV-2 M^{pro} can recognize and cleave the human TRMT1 526-536 sequence located between the TRMT1 methyltransferase and zinc finger domains. M^{pro} is able to cleave this sequence in recombinant full-length TRMT1, as well as in endogenous TRMT1 from human cell lysate. The kinetic parameters determined for M^{pro}-

mediated TRMT1(526-536) peptide proteolysis are very similar to those measured for the known nsp8/9 viral polypeptide cleavage site. We also determined the structure of human TRMT1(526-536) in complex with SARS-CoV-2 M^{pro} C145A, which reveals a distinct binding mode for the TRMT1 substrate peptide that engages the M^{pro} S3′ pocket. Our results show that the human tRNA-modifying enzyme TRMT1 is a viable substrate for the SARS-CoV-2 main protease and provide the structural basis for understanding TRMT1 recognition and proteolysis. Importantly, and concurrent with our own work, Zhang et al report that human TRMT1 is cleaved during SARS-CoV-2 infection and that this leads to loss of m2,2G modification in virally infected cells (36). Together, our studies corroborate and characterize cleavage of human TRMT1 by SARS-CoV-2 M^{pro} from the cellular to the atomic level, and raise important new questions about the roles of tRNA methylation during viral infection.

Viruses have evolved diverse strategies that subvert host protein synthesis in order to downregulate host translation and optimize the translation of viral proteins (37–41). SARS-CoV-2 employs multiple mechanisms to disrupt host protein synthesis (12, 42–47), and M^{pro}-mediated cleavage of human TRMT1 could contribute to the modulation of cellular translation during viral infection. TRMT1 is a tRNA methyltransferase whose m2,2G26 modification activity directly impacts global translational efficiency. Knockouts of TRMT1, or truncations of TRMT1 that remove its zinc finger domain, result in loss of the m2,2G26 tRNA modification and a significant decrease in translation levels in HEK293T cells (21). Our results show that M^{pro}-mediated cleavage of TRMT1 results in the removal of the zinc finger domain, which is expected to substantially weaken TRMT1-tRNA binding affinity and result in loss of enzyme activity (21). Additionally, TRMT1-deficient human cells, or cells with inactivating TRMT1 truncations, exhibit increased sensitivity to oxidative stress and links to neurological dysfunction (48–51). Taken together, the observations that 1) SARS-CoV-2 M^{pro} interacts with TRMT1 in human cells in a catalytic activity-dependent manner (18), 2) full-length TRMT1 is cleaved by M^{pro} in human cell lysate to excise the tRNA-binding zinc finger domain (this work), and 3) loss of the TRMT1 zinc finger domain results in TRMT1 inactivation, decreases in translation, and increased cellular sensitivity to oxidative stress (21), suggest that M^{pro}-mediated cleavage of TRMT1 during SARS-CoV-2 infection may contribute to the downregulation of host translation and oxidative stress-related pathogenesis and phenotypes (52). Future experiments in virally infected cell types will be required to definitively establish how M^{pro}-directed TRMT1 cleavage impacts cellular and viral translation, viral infectivity and propagation, and oxidative stress phenotypes during infection.

TRMT1 displays signatures of rapid evolution in primates, potentially driven by adaptation to ancient viral pathogens, including coronaviruses (53). While rapidly evolving sites can be found near the N- and C-termini of primate TRMT1, we found no evidence of positive selection at the M^{pro}-mediated TRMT1(526-536) cleavage sequence, which is highly conserved (**Figure S5A, Dataset S1**). This evolutionary analysis raises several possibilities regarding the relationship of SARS-CoV and TRMT1. First, it may be that TRMT1(526-536) is targeted for cleavage by M^{pro} during coronavirus infection, but that this TRMT1 motif cannot easily mutate during evolution, due to structural constraints or essential function (54–56), and therefore cannot evade viral proteolysis. Indeed, although the TRMT1(526-536) cleavage sequence is found in a linker region between structured domains, an AlphaFold-predicted (31, 32) TRMT1 structure suggests that many of these residues make contacts with the surface of the TRMT1 MTase domain (**Figure S1**), which may be important for enzyme stability, domain orientation, or function. Second, it may be that TRMT1 cleavage by coronavirus M^{pro} exerts insufficient adaptive pressure on the host. In this case, TRMT1 rapid evolution may be driven by other selective pressure or may be due to intrinsically disordered regions. Third, rapidly evolving sites in primate TRMT1 outside of the cleavage site may reflect evasion of SARS-CoV by providing structural escape through within-protein epistasis (57).

A TRMT1 sequence analysis at the mammalian level further showed that the M^{pro} cleavage site of human TRMT1 (residues 526-536) is highly conserved in most mammals, including in bats that are the reservoir of SARS-CoVs (**Figure 6A, 6B, S5A**). However, one exception is in rodents, where

there has been a Q to K substitution fixed in all Muroidea (**Figure 6C, S5B, Dataset S1**). This TRMT1 Q530K substitution impacts the cleavage motif at the invariant P1 glutamine residue, suggesting it would prevent SARS-CoV M^{pro}-directed TRMT1 cleavage in Muroidea. During the preparation of this manuscript, an independent study has confirmed that human TRMT1 is cleaved by M^{pro}, but that TRMT1 from *Mus musculus* (mouse) and *Mesocricetus auratus* (hamster), which both contain a Q530K substitution, are not cleaved in cells when co-expressed with M^{pro} in HEK293T cells (58). As coronaviruses infect multiple mammalian species, it would be informative to our understanding of SARS-CoV evolution and function to decipher whether Muroidea TRMT1 is naturally resistant to M^{pro}-mediated cleavage, if this resistance confers host advantages or disadvantages during coronavirus infection, and whether this change results from adaptation to ancient coronavirus infection.

Independent of the roles for human TRMT1 and m2,2G RNA modification during SARS-CoV-2 infection, our structural and kinetic results combined with computational analysis highlight distinct features of substrate-protease binding, recognition, and cleavage for SARS-CoV-2 M^{pro}. First, our TRMT1-M^{pro} structure shows an uncommon peptide binding mode in which the P3′ TRMT1 Phe residue is buried in the S3′ pocket on the M^{pro} surface ('P3′-in' conformation). Of the currently known M^{pro}-peptide substrate structures, this unique M^{pro} P3′-in binding mode is only seen for TRMT1 and nsp6/7, though notably the 'P3′-in' conformation has also been observed in a crystal structure of SARS-CoV-1 M^{pro} in complex with its C-terminal auto-processing site (VTFQGKFK), which like TRMT1, contains a phenylalanine at the P3′ site (59). This unique binding mode demonstrates the flexible accommodation of diverse substrate that binding poses in the M^{pro} active site and highlights the availability of the S3′ pocket for inhibitor binding and therapeutic design. Second, consistent with previous mutagenesis studies (33) of M^{pro}, we find that mutations to residues on the surface of M^{pro} involved in direct binding and recognition of substrate, including Met49, Asn142, and Gln189, have little impact on cleavage kinetics. Similarly, the large differences in catalytic efficiency observed for M^{pro}-mediated cleavage of nsp4/5 versus nsp8/9 and TRMT1 are not easily explained by structural analysis of the substrate-bound M^{pro} structures or the M^{pro}-peptide geometries and conformations observed in the MD simulations. Comparing nsp4/5-, nsp8/9-, and TRMT1-bound M^{pro} crystal structures, the scissile amide bond that links substrate residues P1 and P1′ are positioned almost identically in the active site and located at similar distances from the Cys145Ala residue (**Figure S6A**)**.** Likewise, deviations away from 180 degrees in the dihedral angle of the scissile amide bond in these M^{pro}-peptide structures, which could indicate ground state destabilization and help explain accelerated proteolysis for nsp4/5 (60), are also inconsistent with the observed trends in peptide cleavage rates (**Figure S6B**)**.** Molecular dynamics simulations of nsp4/5 and TRMT1 peptide substrates bound to M^{pro} support the observation that the unique P3′-in conformation is favored for TRMT1 binding, but that only subtle differences in geometry are present at the nucleophilic Cys145 residue and scissile peptide bond in the M^{pro}-substrate complexes, which cannot account for the order of magnitude differences in observed cleavage kinetics between ns4/5 and TRMT1 substrates. These observations suggest that kinetic discrimination of substrate cleavage happens during later steps of the M^{pro}-mediated proteolysis reaction.

Our work defines the structural basis for understanding recognition and cleavage of the human tRNA methyltransferase TRMT1 by the SARS-CoV-2 main protease. Zhang et al provide evidence that TRMT1 is proteolyzed and inactivated during SARS-CoV-2 infection, which would be predicted to affect host and viral translation and increase cellular sensitivity to redox stress (36). TRMT1 cleavage may therefore have important implications for understanding the ability of SARS-CoV-2 to hijack host protein synthesis and impact cellular phenotypes and COVID-19 pathogenesis. Additionally, our TRMT1-M^{pro} structure highlights a distinct substrate binding mode that reveals the S3′ pocket on the surface of the main protease. SARS-Cov-2 M^{pro} is vital for viral propagation and is thus a major target for antiviral drug development (61–65). Consideration of the often-hidden S3′ pocket on the M^{pro} surface could aid in the design of potent and specific SARS-CoV-2 M^{pro} inhibitors. Currently, the vast majority of inhibitors target the S2 and S1 sites of M^{pro}, with peptidyl

inhibitors designed with glutamine and leucine analogs for the P2 and P1 positions, respectively (66). Our structural results suggest the flexible S3′ pocket could be utilized in future therapeutic design, and together with our kinetic analysis adds to the growing understanding of M^{pro} substrate recognition and cleavage efficiency that is critical for potent and specific targeting of M^{pro} to prevent or treat infection.

Materials and Methods

Cloning, Protein Expression, and Purification

M^{pro} constructs used for biochemical and kinetic studies were subcloned into a pETARA expression vector containing N-terminal GST and C-terminal his tags using Gibson assembly. Human codon optimized wild-type M^{pro} was obtained from Addgene (catalog #141370). M^{pro} single point mutants (M49A, N142A, Q189A) were introduced using site-directed mutagenesis by whole plasmid PCR. Each construct contained the M^{pro} autocleavage sequence (AVLQ) after the N-terminal GST tag to allow for self-cleavage and a native M^{pro} N-terminus. Constructs were transformed into *E. coli* Rosetta™(DE3)pLysS cells and plated on LB agar plates with 50 µg/ml ampicillin. Overnight seed cultures were grown from single colonies in LB media with 50 µg/ml ampicillin, at 37°C with shaking at 200 rpm. A 1:100 dilution of seed culture to 1L LB media was prepared and grown at 37°C with shaking at 200 rpm. After an OD_{600} of 0.6 was reached, cells were induced with 1mM IPTG and incubated overnight at 18°C with shaking. Cultures were centrifuged at 7,500 x g and the supernatant was discarded. Harvested cells were sonicated in lysis buffer (25mM Tris, 300mM NaCl, 10mM Imidazole pH 8.0), centrifuged at 14,500 x g for 45 minutes, and recovered, clarified lysate was loaded onto 1.5 mL (bed volume) equilibrated Thermo Scientific HisPur Ni-NTA Resin for 30 minutes with gentle mixing. Resin was added to a standard gravity column and washed with two 25mL washes of lysis buffer and two 25mL washes with lysis buffer + 25mM imidazole, and eluted with 10 mL lysis buffer + 250mM imidazole. Eluate was concentrated to ~2mL using a using 10kDa MWCO centrifugal concentrator and applied to a Cytiva HiLoad 16/600 Superdex 200 pg column for size exclusion chromatography. M^{pro}-containing fractions were concentrated to between 10-25 mg/mL, flash frozen with liquid nitrogen, and stored at -70°C in 50 mM Tris, 1 mM EDTA, 2 mM DTT, pH 7.3.

Catalytically inactive M^{pro} C145A mutants used for protein crystallography experiments were prepared by subcloning M^{pro} C145A from Addgene (catalog #141371) into a pET28a expression vector containing an N-terminal His-tag and tobacco etch virus (TEV) protease cleavage site (ENLYFQGS). The M^{pro} C145A construct was transformed and expressed in *E. coli* as described above. Cultures were centrifuged at 7,500 x g and the supernatant was discarded. Harvested pellets were sonicated in lysis buffer (25mM Tris, 300mM NaCl, pH 8.0) and the clarified the lysate by centrifugation at 14, 500 x g. Recovered lysate was loaded onto 3.75 mL (bed volume) equilibrated Thermo Scientific HisPur Ni-NTA Resin for 30 minutes with gentle mixing. Resin was added to a standard gravity column and washed with two 25mL washes of lysis buffer and two 25mL washes with lysis buffer + 25mM Imidazole, and eluted with lysis buffer + 300mM Imidazole. Eluate was incubated overnight at room temperature with recombinant purified TEV protease at a 1:50 molar ratio to cleave the N-terminal His-tag. A Ni-NTA back pass was performed and the flow-through was concentrated, buffer exchanged into 25mM Tris, 25mM NaCl, pH 7.5 using a using 10kDa MWCO centrifugal concentrator, and applied to a Cytiva HiTrap 5mL Q HP column. Protein was eluted with a gradient of 25 mM to 1M NaCl and M^{pro}-containing fractions were pooled, concentrated and applied to a Cytiva HiLoad 16/600 Superdex 200 pg size exclusion chromatography column and eluted with an isocratic gradient using 25mM Tris 25mM NaCl pH 7.5. The protein-containing fractions were concentrated to approximately 26 mg/mL, flash frozen with 10% glycerol in liquid nitrogen, and stored at -70°C.

Full-length human TRMT1 expression constructs used for proteolysis experiments were generated using a codon optimized gBlock obtained from IDT containing the human full-length TRMT1

sequence which was subcloned into a pET28a expression vector with an N-terminal His-tag and TEV protease cleavage site (ENLYFQGS). The TRMT1 construct was transformed into Rosetta™(DE3)pLysS *E. coli* as described above. Harvested cells were sonicated in lysis buffer (PBS, cOmplete EDTA-free Protease Inhibitor Cocktail, 0.5mg/mL lysozyme), clarified at 14,500 x g, and recovered lysate was loaded onto a pre-equilibrated Cytiva 5mL HisTrap column. The loaded HisTrap column washed with lysis buffer + 25mM Imidazole, and protein was eluted using a gradient elution with lysis buffer + 300mM Imidazole. Fractions containing TRMT1 were identified by Western blotting and were pooled, concentrated, buffer exchanged into 20 mM HEPES pH 7.0, 150 mM NaCl, flash frozen in liquid nitrogen, and stored at -70°C.

Crystallography and Structure Determination

Purified M^{pro} C145A was incubated at room temperature for 1 hour with TRMT1(526-536) peptide (EPRLQANFTIR, synthetic peptide obtained from Peptide 2.0) in 25mM Tris, 25mM NaCl, pH 7.4at a 1:3 molar ratio with the final concentrations of 7 mg/mL M^{pro} C145A and 619 μM TRMT1(526-536). The M^{pro} and TRMT1 solution was mixed 1:1 with well solutions consisting of 20-21.5% PEG 3350 and 100mM NaCl in 24-well hanging drop VDX plates with a final drop volume of 1 μL. Initial crystals were harvested, crushed, and seeded into new hanging drops, with components as listed above, using cat whiskers. Single crystals were harvested and flash frozen in liquid nitrogen with 20% glycerol in well solution as a cryoprotectant. Diffraction data were collected at the National Synchrotron Light Source II (NSLS II) Highly Automated Macromolecular Crystallography (AMX) (67) beamline 17-ID-1 at the Brookhaven National Laboratory on an Eiger 9M Pixel detector at 100 K and a wavelength of 0.920219 Å. Diffraction data were indexed, integrated, and scaled using XDS. The M^{pro}-TRMT1 structure was solved in space group $P2_12_12_1$ using the Phaser package in the CCP4 suite and a modified PDB 7MGS as search model. After initial rounds of refinement in PHENIX to model M^{pro} residues, the TRMT1 peptide residues 526 – 534 were manually built into the F_o-F_c map using COOT. Subsequent rounds of automated refinement and water placement using PHENIX and manual adjustments including modeling sodium ions and glycerol molecules in COOT were used to obtain the final structure.

Peptide Cleavage Kinetic Assays

Fluorogenic assays measuring kinetic parameters for cleavage of peptide substrates were carried out in triplicate in Corning Low Volume 384-well Black Flat Bottom Polystyrene NBS Microplates following a similar procedure as Lee, *et al*. (68). Peptide cleavage reactions were carried out at 50 nM M^{pro} enzyme in 50mM Tris pH 7.3, 1mM EDTA, 2mM DTT, and 20% DMSO, with quenched fluorescent peptide substrates ranging in concentration from 0.097 – 100 μM. Synthetic peptide substrates were obtained from Peptide 2.0: TRMT1 (MCA-EPRLQANFTIR-K(Dnp)K) or nsp4/5 (MCA-SAVLQSGFRKM-K(Dnp)K), where MCA = 7-Methoxycoumarin-4-acetic acid and Dnp = dinitrophenyl. Using a Tecan Spark microplate reader, the fluorescence intensity was monitored every 10 seconds over a 3-minute time course, with excitation at 320 nm and emission at 405 nm. A calibration curve of MCA-AVLQ product fluorescence intensities from 12 to 0.006 μM was measured and used to generate an RFU to μM conversion factor. A correction for the inner filter effect (IFE_{corr}) was determined using the formula IFE_{corr} = [$fluorescence_{MCA_product+peptide}$ – $fluorescence_{peptide}$] / $fluorescence_{MCA}$ where $fluorescence_{MCA_product+peptide}$ = the fluorescence of MCA-AVLQ mixed with quenched peptide substrate, $fluorescence_{peptide}$ = the fluorescence of the peptide alone, $fluorescence_{MCA}$ = fluorescence of MCA-AVLQ product alone (**Dataset S2**) (69). Plots of initial rate (μM/min) versus peptide substrate concentration were fit to the Michaelis-Menten equation to determine V_{max} and K_M kinetic parameters for M^{pro}-mediated peptide cleavage using Origin 2021 software.

Mammalian Cell Culture

Freshly passaged HEK293T cells were gifted from Prof. Catherine Grimes (University of Delaware). Cells were incubated at 37°C with 5% CO_2 in growth media (DMEM media with 10% FBS and

Penicillin (100 U/mL)/Streptomycin (100 µg/mL)). Growth media was replaced every 3 days, and cells were trypsinized and passaged approximately every 7 days, when cells reached 80-90% confluency. To passage, growth media was aspirated, and cells were gently washed with prewarmed PBS. PBS was then aspirated and a 1:2.5 mixture of 0.25% Trypsin in HBSS with 0.2 g/L EDTA to PBS was added to adherent cells and incubated at 37°C for 2 minutes to disaggregate. A 3X dilution with growth media was used to inactivate trypsin. The cells were gently mixed to attain a homogenous suspension and diluted 1:20 with growth media and added into a new cell culture plate. For lysis, cells were trypsinized as described above, and the resulting cell suspension was spun down at 1250 RPM for 3 minutes. Cell pellets were washed in ice cold PBS 3 times, and spun at 1250 RPM for 3 minutes, aspirating PBS in between each wash. Lysis was performed on a PBS suspension of cells by performing 3 consecutive freeze/thaw cycles by flash freezing in liquid nitrogen and thawing at 37°C. Lysate was collected after centrifugation at 13,000 RPM for 10 minutes and stored at -70°C.

Proteolysis Assay

TRMT1 proteolysis reactions with recombinant TRMT1 isolated from *E. coli* were performed with 10 µM Wild-type or C145A M^{pro} in 20mM HEPES,105mM NaCl, 20% glycerol, pH 7.0. Recombinant full-length TRMT1 from *E. coli* was diluted 7-fold. Reaction was incubated at 37°C and time points were quenched by adding to SDS-PAGE sample buffer (50 mM Tris, pH 6.8, 1.0% (w/v) SDS, 10% (v/v) Glycerol, 0.1% (w/v) Bromophenol blue, 0.1 M DTT) and boiled for 5 minutes. TRMT1 levels and fragment sizes at different reaction timepoints were assessed by Western blot.

TRMT1 proteolysis assays using endogenous human TRMT1 from HEK293T lysate were carried out using cell lysate preincubated with 1mM phenylmethylsulfonyl fluoride (PMSF) prior to reaction to prevent proteolysis by mammalian-specific proteases. M^{pro} proteolysis reactions were performed with 10 µM Wild-type or C145A M^{pro} in 25mM Tris, 25mM NaCl, 20% glycerol, pH 7.5. Reaction was incubated at 37°C and time points were quenched by adding to SDS-PAGE sample buffer (50 mM Tris, pH 6.8, 1.0% (w/v) SDS, 10% (v/v) Glycerol, 0.1% (w/v) Bromophenol blue, 0.1 M DTT) and boiled for 5 minutes. TRMT1 levels and fragment sizes at different reaction timepoints were assessed by Western blot.

Western Blot

Quenched samples from the proteolysis assays were loaded (10uL) onto Bio-Rad TGX 4-15% Polyacrylamide gels and run for 30 minutes at 180V. Gels were blotted onto PVDF membranes using a Bio-Rad Trans blot Turbo for 7 minutes at 2.5 Amps. The blot was incubated in blocking solution (5% non-fat milk in 1X Tris-Buffer Saline with 0.1% Tween) at room temperature for 1 hour. All antibodies were diluted in blocking solution. TRMT1-specific antibodies corresponding to amino acid regions 460-659 (Invitrogen Rabbit Anti-TRMT1 (ref: PA5-96585)) and 609-659 (Bethyl Laboratories Rabbit Anti-TRMT1 (ref: A304-205A)) were utilized for primary antibody staining at 1:2,000 dilution and overnight incubation at 4°C. Extensive washing of blot with 1X Tris-buffer Saline with 0.1% Tween were performed after primary antibody staining. Invitrogen Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP (ref: A16096) was used for secondary staining at a dilution of 1:10,000 for 1 hour at room temperature. Extensive washing of blot with 1X Tris-buffer Saline with 0.1% Tween were performed after secondary antibody staining. Clarity™ Western ECL Substrate (ref: 1705060) was added to blot and incubated at room temperature for 5 minutes. Western blots were visualized on a Protein Simple FluorChem R imager.

Molecular Dynamics Simulation

Extensive molecular dynamics simulations in explicit water have been carried out to investigate M^{pro}-nsp4/5 and M^{pro}-TRMT1 complexes starting from their corresponding crystal structures (PDB id: 7MGS (15) and 8D35), respectively. In original crystal structures, the catalytic Cys145 has been mutated to Ala, M^{pro}-nsp4/5 is in the monomer form, while the M^{pro}-TRMT1 complex is in the dimer

form. Since the active form of M^{pro} should be dimer (70), the structure of the M^{pro}-nsp4/5 (PDB id: 7MGS) was superimposed to the M^{pro}-TRMT1 (PDB id: 8D35) to model the dimer form of M^{pro}-nsp4/5, and the catalytic Cys145 in both complexes has been modeled using 'swapaa' command in ChimeraX (71). The original crystal water molecules and ions within 5 Å of the protein-substrate complex are kept. PDB2PQR (72) was employed to add hydrogen atoms to both complexes at pH 7. Subsequently, the His41 was manually changed to Hip41(positively charged) by adding an additional hydrogen atom to the imidazole ring in ChimeraX, and Cys145 was manually changed to Cym(negatively charged) by removing the hydrogen atom on the thiol group in ChimeraX. Each system was neutralized by adding counterions for Na^+ and Cl^-, and solvated in a rectangular water box of TIP3P water molecules with 12 Å buffer.

For each prepared simulation system, 3000 steps of steepest descent plus 2000 steps of conjugate minimization was performed with harmonic restraints using a force constant of 20 $kcal \cdot mol^{-1} \cdot Å^{-2}$ applied to all heavy atoms coming from the crystal structure, and then the whole system was minimized without restraints by 7000 steps of steepest descent and 3000 steps of conjugate gradient minimization. After minimization, three independent equilibration and MD replicas were carried out for 100ns with the same starting structural configuration but different initial velocities. The equilibration was conducted in five steps: (i) 50 ps constant volume ensemble(NVT) MD simulation with 10 $kcal \cdot mol^{-1} \cdot Å^{-2}$ restraints on all heavy atoms from the crystal, and the whole system was heated from 10 to 300K gradually. (ii) 50 ps isothermal isobaric ensemble (NPT) MD with 10 $kcal \cdot mol^{-1} \cdot Å^{-2}$ restraints on all heavy atoms from the crystal at 300 K. (iii) 200 ps NPT MD with 5 $kcal \cdot mol^{-1} \cdot Å^{-2}$ restraints on all heavy atoms from the crystal at 300 K. (iv) 200 ps NPT MD with 2 $kcal \cdot mol^{-1} \cdot Å^{-2}$ restraints on all heavy atoms from the crystal at 300 K. (v) 200 ps NPT MD with 1 $kcal \cdot mol^{-1} \cdot Å^{-2}$ restraints on all heavy atoms from the crystal at 300 K. Finally, production MD simulations were carried out for 100 ns at a constant temperature of 300 K and a constant pressure of 1 atm. Langevin thermostat and Berendsen barostat methods were employed to maintain the temperature and pressure, respectively. The atomic coordinates of the complexes were saved every 1ps to obtain the trajectories for analysis.

All MD simulations were conducted with Amber 20 (73) package using Amber FF 14SB force field (74). AmberTools (73) was utilized for preparing topology and coordinate files for the simulated systemsA cutoff of 10 Å was set for calculating van der Waals interactions, and the particle mesh Ewald (PME) (75) method with a cutoff of 10 Å was employed to treat electrostatic interactions. The SHAKE algorithm (76) was used to constrain covalent bonds to allow the integration time step of 2 fs.

Evolutionary analysis of mammalian TRMT1 orthologs

For the mammalian-wide sequence analyses, the TRMT1 amino acid alignment was retrieved with OrthoMaM (Orthologous Mammalian Markers) v10c (77), using human TRMT1 (ENSG00000104907) as query. The cleavage site was located and the sequence logo of the 526-536 region was generated using WebLogo3 (https://weblogo.threeplusone.com/).

For the primate phylogenetic analyses, TRMT1 orthologous sequences were retrieved using the DGINN pipeline with the human TRMT1 CCDS12293 as query (78). TRMT1 sequences from additional primate species were retrieved using NCBI Blastn (Dataset S1).
For the rodent phylogenetic analyses, the rodentia orthologous protein and mRNA reference sequences to the human TRMT1 (ENSG00000104907) gene were collected from OrthoMaM and ncbi HomoloGene (https://www.ncbi.nlm.nih.gov/homologene) (Dataset S1). Because most rodent species lack a reference CDS, we used BLOSUM62 matrix implemented in GeneWise from EMBL-EBI tools (79, 80), to identify the TRMT1 ORFs. Codon alignments of primate and rodent TRMT1s were performed using WebPrank (81) with default settings for primates (trust insertions +F, gap rate= 0.05, gap length= 5, K= 2) and with a gap rate of 0.1 for rodents. Positive selection analyses on the codon alignments were performed with HYPHY/Datamonkey (82, 83), using two branch-site

models, BUSTED (version 3.1)(84) and aBSREL (version 2.2) (85), and two site- specific models, MEME (version 2.1.2)(86), FUBAR (version 2.2) (87).

To identify missense polymorphisms or variants in TRMT1 at minor allele frequency above 0.005 in human population, we mined the dbSNP database https://www.ncbi.nlm.nih.gov/snp/ .

Acknowledgments

We would like to thank A. Soares at Brookhaven National Laboratory, National Synchrotron Light Source II and the Schmitz lab at University of Delaware for assistance with X-ray data collection and X. Liu at UCSF for advice on structure refinement. We are grateful to J.S.M.'s cat Percival for the generous donation of whiskers used in protein crystal seeding experiments. We thank A. Cimarelli and the LP2L team members at CIRI for feedback.

This work was supported by the US National Institutes of Health, National Institute of General Medical Sciences, under awards R35 GM143000 to J.S.M., R35 GM127040 to Y.Z., T32 GM133395 CBI fellowship to A.D., and P20 GM104316. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The work was further supported by the CNRS and the French Agence Nationale de la Recherche (ANR), under grant ANR-20-CE15-0020-01 (project "BATantiVIR") and by a grant from the French Research Agency on HIV and Emerging Infectious Diseases ANRS/MIE (#ECTZ118944) to L.E.

Computational resources were provided by NYU-ITS.

This research used resources at AMX beamline 17-ID-1 of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. The Center for BioMolecular Structure (CBMS) is primarily supported by the National Institutes of Health, National Institute of General Medical Sciences (NIGMS) through a Center Core P30 Grant (P30GM133893), and by the DOE Office of Biological and Environmental Research (KP1607011).

Data availability

Coordinates and structure factors were deposited in the Protein Data Bank with accession code PDB 8D35. Data for sequence analyses are further available in Dataset S1, and https://doi.org/10.6084/m9.figshare.22004474 and https://doi.org/10.6084/m9.figshare.22004492. All other data are available from the corresponding author upon reasonable request

References

1. G. W. H. Organization, WHO COVID-19 Dashboard. **2023** (2023).

2. C. for S. Science, E. (CSSE) at J. H. University (JHU), COVID-19 Dashboard. **2023** (2023).

3. A. B. Suthar, *et al.*, Public health impact of covid-19 vaccines in the US: observational study. *BMJ Online* **377**, e069317 (2022).

4. K. Rahmani, *et al.*, The effectiveness of COVID-19 vaccines in reducing the incidence, hospitalization, and mortality from COVID-19: A systematic review and meta-analysis. *Front. Public Health* **10**, 873596 (2022).

5. O. J. Watson, *et al.*, Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *Lancet Infect. Dis.* **22**, 1293–1302 (2022).

6. I. Mohammed, *et al.*, The efficacy and effectiveness of the COVID-19 vaccines in reducing infection, severity, hospitalization, and mortality: a systematic review. *Hum. Vaccines Immunother.* **18**, 2027160 (2022).

7. S. K. Burley, *et al.*, RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.* **49**, D437–D451 (2021).

8. M. Wilamowski, *et al.*, Transient and stabilized complexes of Nsp7, Nsp8, and Nsp12 in SARS-CoV-2 replication. *Biophys. J.* **120**, 3152–3165 (2021).

9. L. Yan, *et al.*, Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nat. Commun.* **11**, 5874 (2020).

10. D. Shin, *et al.*, Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nat. Lond.* **587**, 657–662 (2020).

11. Y.-X. Zheng, *et al.*, Nsp2 has the potential to be a drug target revealed by global identification of SARS-CoV-2 Nsp2-interacting proteins. *Acta Biochim. Biophys. Sin.* **53**, 1134–1141 (2021).

12. M. Thoms, *et al.*, Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* **369**, 1249–1255 (2020).

13. S. Yuan, *et al.*, Nonstructural Protein 1 of SARS-CoV-2 Is a Potent Pathogenicity Factor Redirecting Host Protein Synthesis Machinery toward Viral RNA. *Mol. Cell* **80**, 1055-1066.e6 (2020).

14. A. M. Shaqra, *et al.*, Defining the Substrate Envelope of SARS-CoV-2 Main Protease to Predict and Avoid Drug Resistance. *bioRxiv*, 2022.01.25.477757 (2022).

15. E. A. MacDonald, *et al.*, Recognition of Divergent Viral Substrates by the SARS-CoV-2 Main Protease. *ACS Infect. Dis.* **7**, 2591–2595 (2021).

16. Y. Zhao, *et al.*, Structural basis for replicase polyprotein cleavage and substrate specificity of main protease from SARS-CoV-2. *Proc. Natl. Acad. Sci.* **119**, e2117142119 (2022).

17. M. A. Hameedi, *et al.*, Structural and functional characterization of NEMO cleavage by SARS-CoV-2 3CLpro. *Nat. Commun.* **13**, 5285 (2022).

18. D. E. Gordon, *et al.*, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nat. Lond.* **583**, 459–468 (2020).

19. M. Bouhaddou, *et al.*, Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**, 1181 (2020).

20. H.-J. Chou, E. Donnard, H. T. Gustafsson, M. Garber, O. J. Rando, Transcriptome-wide Analysis of Roles for tRNA Modifications in Translational Regulation. *Mol. Cell* **68**, 978-992.e4 (2017).

21. J. M. Dewe, B. L. Fuller, J. M. Lentini, S. M. Kellner, D. Fu, TRMT1-Catalyzed tRNA modifications are required for redox homeostasis to ensure proper cellular proliferation and oxidative stress survival. *Mol. Cell. Biol.* **37** (2017).

22. M. Miczi, *et al.*, Identification of Host Cellular Protein Substrates of SARS-COV-2 Main Protease. *Int. J. Mol. Sci.* **21**, 9523 (2020).

23. T. Koudelka, *et al.*, N-Terminomics for the Identification of In Vitro Substrates and Cleavage Site Specificity of the SARS-CoV-2 Main Protease. *PROTEOMICS* **21**, e2000246-n/a (2021).

24. M. Moustaqil, *et al.*, SARS-CoV-2 proteases PLpro and 3CLpro cleave IRF3 and critical modulators of inflammatory pathways (NLRP12 and TAB1): implications for disease presentation across species. *Emerg. Microbes Infect.* **10**, 178–195 (2021).

25. N. Yucel, *et al.*, Prediction and validation of host cleavage targets of SARS-CoV-2 3C-like protease. *bioRxiv*, 2022.01.17.476677 (2022).

26. S. Zhang, J. Wang, G. Cheng, Protease cleavage of RNF20 facilitates coronavirus replication via stabilization of SREBP1. *Proc. Natl. Acad. Sci. - PNAS* **118**, 1 (2021).

27. I. Pablos, *et al.*, Mechanistic insights into COVID-19 by global analysis of the SARS-CoV-2 3CLpro substrate degradome. *Cell Rep. Camb.* **37**, 109892 (2021).

28. B. Meyer, *et al.*, Characterising proteolysis during SARS-CoV-2 infection identifies viral cleavage sites and cellular targets with therapeutic potential. *Nat. Commun.* **12**, 5553 (2021).

29. K. Nakagawa, K. G. Lokugamage, S. Makino, Viral and Cellular mRNA Translation in Coronavirus-Infected Cells. *Adv. Virus Res.* **96**, 165–192 (2016).

30. R. Gorkhali, *et al.*, Structure and Function of Major SARS-CoV-2 and SARS-CoV Proteins. *Bioinforma. Biol. Insights* **15**, 11779322211025876 (2021).

31. M. Varadi, *et al.*, AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).

32. J. Jumper, *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

33. J. M. Flynn, *et al.*, Comprehensive fitness landscape of SARS-CoV-2 Mpro reveals insights into viral resistance mechanisms. *bioRxiv*, 2022.01.26.477860 (2022).

34. E. S. Radisky, J. K. D. E, A Clogged Gutter Mechanism for Protease Inhibitors. *Proc. Natl. Acad. Sci. - PNAS* **99**, 10316–10321 (2002).

35. D. E. Gordon, *et al.*, Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* (2020) https:/doi.org/10.3929/ethz-b-000531946.

36. K. Zhang, *et al.*, Proteolytic cleavage and inactivation of the TRMT1 tRNA modification enzyme by SARS-CoV-2 main protease. *bioRxiv* (2023) https:/doi.org/10.1101/2023.02.10.527147.

37. S. de Breyne, *et al.*, Translational control of coronaviruses. *Nucleic Acids Res.* **48**, 12502–12522 (2020).

38. E. Jan, I. Mohr, D. Walsh, A Cap-to-Tail Guide to mRNA Translation Strategies in Virus-Infected Cells. *Annu. Rev. Virol.* **3**, 283–307 (2016).

39. A. Sharma, A. Yilmaz, K. Marsh, A. Cochrane, K. Boris-Lawrie, Thriving under Stress: Selective Translation of HIV-1 Structural Protein mRNA during Vpr-Mediated Impairment of eIF4E Translation Activity. *PLoS Pathog.* **8**, e1002612 (2012).

40. D. Walsh, I. Mohr, Viral subversion of the host protein synthesis machinery. *Nat. Rev. Microbiol.* **9**, 860–875 (2011).

41. L. O. Roberts, C. L. Jopling, R. J. Jackson, A. E. Willis, "Chapter 9 Viral Strategies to Subvert the Mammalian Translation Machinery" in Progress in Molecular Biology and Translational Science., (Elsevier Inc, 2009), pp. 313–367.

42. Y. Finkel, *et al.*, SARS-CoV-2 uses a multipronged strategy to impede host protein synthesis. *Nat. Lond.* **594**, 240–245 (2021).

43. J. C.-C. Hsu, M. Laurent-Rolle, J. B. Pawlak, C. B. Wilen, P. Cresswell, Translational shutdown and evasion of the innate immune response by SARS-CoV-2 NSP14 protein. *Proc. Natl. Acad. Sci. - PNAS* **118**, 1 (2021).

44. M. Puray-Chavez, *et al.*, The Translational Landscape of SARS-CoV-2-infected Cells Reveals Suppression of Innate Immune Genes. *mBio* **13**, e0081522 (2022).

45. G. Eriani, F. Martin, Viral and cellular translation during SARS-CoV-2 infection. *FEBS Open Bio* **12**, 1584–1601 (2022).

46. D. Zhang, L. Zhu, Y. Wang, P. Li, Y. Gao, Translational Control of COVID-19 and Its Therapeutic Implication. *Front. Immunol.* **13**, 857490 (2022).

47. A. K. Banerjee, *et al.*, SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. *Cell* **183**, 1325-1339.e21 (2020).

48. K. Blaesius, *et al.*, Mutations in the tRNA methyltransferase 1 gene TRMT1 cause congenital microcephaly, isolated inferior vermian hypoplasia and cystic leukomalacia in addition to intellectual disability. *Am. J. Med. Genet. A.* **176**, 2517–2521 (2018).

49. B. Davarniya, *et al.*, The Role of a Novel TRMT1 Gene Mutation and Rare GRM1 Gene Defect in Intellectual Disability in Two Azeri Families. *PloS One* **10**, e0129631 (2015).

50. H. Najmabadi, *et al.*, Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nat. Lond.* **478**, 57–63 (2011).

51. K. Zhang, *et al.*, An intellectual disability-associated missense variant in TRMT1 impairs tRNA modification and reconstitution of enzymatic activity. *Hum. Mutat.* **41**, 600–607 (2020).

52. R. Cecchini, A. L. Cecchini, SARS-CoV-2 infection pathogenesis is related to oxidative stress as a response to aggression. *Med. Hypotheses* **143**, 110102 (2020).

53. M. Cariou, *et al.*, Distinct evolutionary trajectories of SARS-CoV-2-interacting proteins in bats and primates identify important host determinants of COVID-19. *Proc. Natl. Acad. Sci.* **119**, e2206610119 (2022).

54. B. Murrell, T. Vollbrecht, J. Guatelli, J. O. Wertheim, The Evolutionary Histories of Antiretroviral Proteins SERINC3 and SERINC5 Do Not Support an Evolutionary Arms Race in Primates. *J. Virol.* **90**, 8085–8089 (2016).

55. F. Abdul, *et al.*, Smc5/6 Antagonism by HBx Is an Evolutionarily Conserved Function of Hepatitis B Virus Infection in Mammals. *J. Virol.* **92**, e00769-18 (2018).

56. D. Enard, L. Cai, C. Gwennap, D. A. Petrov, Viruses are a dominant driver of protein adaptation in mammals. *eLife* **5**, e12469 (2016).

57. T. N. Starr, J. W. Thornton, Epistasis in protein evolution. *Protein Sci. Publ. Protein Soc.* **25**, 1204–1218 (2016).

58. J.-L. Lu, X.-L. Zhou, SARS-CoV-2 main protease Nsp5 cleaves and inactivates human tRNA methyltransferase TRMT1. *J. Mol. Cell Biol.*, mjad024 (2023).

59. T. Muramatsu, *et al.*, SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proc. Natl. Acad. Sci.* **113**, 12997–13002 (2016).

60. Y. Zhou, D. Xie, Y. Zhang, Amide Rotation Hindrance Predicts Proteolytic Resistance of Cystine-Knot Peptides. *J. Phys. Chem. Lett.* **7**, 1138–1142 (2016).

61. A. M. Burkhardt, A. Lu, I. Asante, S. Louie, “Pharmacologic Therapeutics for COVID-19” in Coronavirus Disease 2019 (COVID-19)., (2023), pp. 290–318.

62. N. Ćwilichowska, K. W. Świderska, A. Dobrzyń, M. Drąg, M. Poręba, Diagnostic and therapeutic potential of protease inhibition. *Mol. Aspects Med.* **88**, 101144 (2022).

63. S. G. Katre, *et al.*, Review on development of potential inhibitors of SARS-CoV-2 main protease (MPro). *Future J. Pharm. Sci.* **8**, 36 (2022).

64. K. Steuten, *et al.*, Challenges for Targeting SARS-CoV-2 Proteases as a Therapeutic Strategy for COVID-19. *ACS Infect. Dis.* **7**, 1457–1468 (2021).

65. A. Narayanan, *et al.*, Identification of SARS-CoV-2 inhibitors targeting Mpro and PLpro using in-cell-protease assay. *Commun. Biol.* **5**, 169 (2022).

66. L. Agost-Beltrán, *et al.*, Advances in the Development of SARS-CoV-2 Mpro Inhibitors. *Mol. Basel Switz.* **27**, 2523 (2022).

67. D. K. Schneider, *et al.*, AMX – the highly automated macromolecular crystallography (17-ID-1) beamline at the NSLS-II. *J. Synchrotron Radiat.* **29**, 1480–1494 (2022).

68. J. Lee, *et al.*, Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. *Nat. Commun.* **11**, 5877 (2020).

69. Y. Liu, *et al.*, Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. *Anal. Biochem.* **267**, 331–335 (1999).

70. L. Zhang, *et al.*, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **368**, 409–412 (2020).

71. E. F. Pettersen, *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **30**, 70–82 (2021).

72. E. Jurrus, *et al.*, Improvements to the APBS biomolecular solvation software suite. *Protein Sci. Publ. Protein Soc.* **27**, 112–128 (2018).

73. D. Case, *et al.*, AMBER 2020: University of California. *San Franc.* (2020).

74. J. A. Maier, *et al.*, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

75. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

76. H. C. Andersen, Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **52**, 24–34 (1983).

77. C. Scornavacca, *et al.*, OrthoMaM v10: Scaling-Up Orthologous Coding Sequence and Exon Alignments with More than One Hundred Mammalian Genomes. *Mol. Biol. Evol.* **36**, 861–862 (2019).

78. L. Picard, *et al.*, DGINN, an automated and highly-flexible pipeline for the detection of genetic innovations on protein-coding genes. *Nucleic Acids Res.* **48**, e103 (2020).

79. E. Birney, M. Clamp, R. Durbin, GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004).

80. F. Madeira, *et al.*, The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).

81. A. Löytynoja, N. Goldman, webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* **11**, 579 (2010).

82. S. L. Kosakovsky Pond, *et al.*, HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol. Biol. Evol.* **37**, 295–299 (2020).

83. S. Weaver, *et al.*, Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol. Biol. Evol.* **35**, 773–777 (2018).

84. B. Murrell, *et al.*, Gene-wide identification of episodic selection. *Mol. Biol. Evol.* **32**, 1365–1371 (2015).

85. M. D. Smith, *et al.*, Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).

86. B. Murrell, *et al.*, Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8**, e1002764 (2012).

87. B. Murrell, *et al.*, FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol. Biol. Evol.* **30**, 1196–1205 (2013).

Figures and Tables

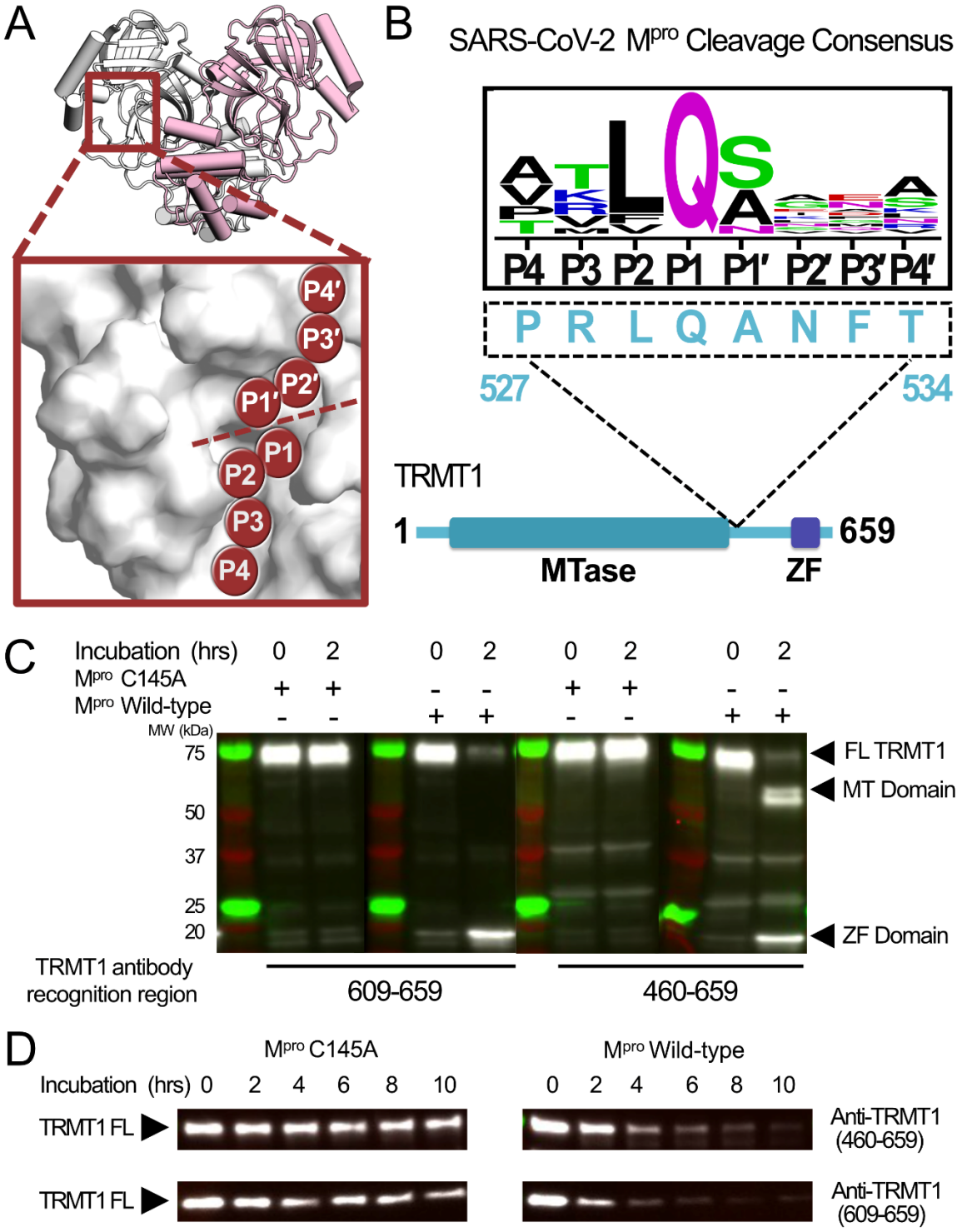

Figure 1. SARS-CoV-2 M^{pro} cleaves full-length human TRMT1. **A)** Overview of the structure of the SARS-CoV-2 M^{pro} homodimer (PDB 7BB2) with substrate peptide residues (P4-P3-P2-P1-P1′-P2′-P3′-P4′) illustrated in the M^{pro} active site (inset); proteolytic cleavage takes place between substrate

residues P1 and P1′ (dotted line). **B)** The TRMT1(527-534) sequence found in a linker region between the TRMT1 SAM methyltransferase (MTase) and Zinc Finger (ZF) domains is consistent with the SARS-CoV-2 M^{pro} cleavage consensus sequence. **C)** Western blots of recombinantly expressed full-length TRMT1 incubated with 10 μM catalytically inactive (Cys145Ala) or active (Wild-type) SARS-CoV-2 M^{pro} at 37°C. Incubation with WT M^{pro} results in proteolysis of FL TRMT1 and the appearance of cleavage products corresponding the ZF domain (observed with both anti-TRMT1(609-659) and anti-TRMT1(460-659) antibodies) and the MTase domain (observed with only anti-TRMT1(460-659) antibody). **D)** Western blots of endogenous human TRMT1 in HEK293T cell lysate incubated with 10 μM of either catalytically inactive (Cys145Ala) or active (Wild-type) M^{pro} at 37°C. Endogenous FL TRMT1 is stable in human cell lysate over the course of a 10-hour incubation with C145A M^{pro} (left) and is rapidly proteolyzed upon incubation with WT M^{pro} (right).

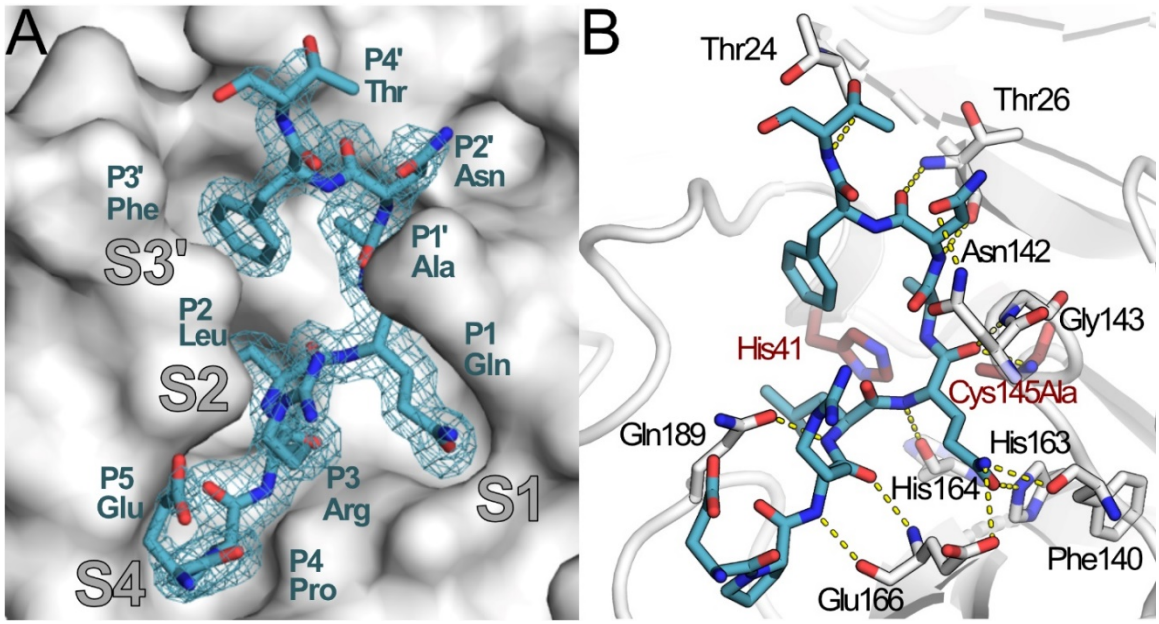

Figure 2. Structure of human TRMT1(526-536) peptide bound to SARS-CoV-2 M^{pro}. **A)** TRMT1 peptide bound in M^{pro} active site showing substrate binding pockets S1, S2, S4, and S3′. F_o-F_c omit electron density map of TRMT1 peptide bound to M^{pro} contoured at 2σ. TRMT1 Gln P1, an ultra-conserved residue in M^{pro} cleavage consensus which is critical for M^{pro}-mediated proteolysis, is nestled in the S1 subsite pocket of the M^{pro} active site. **B)** Direct hydrogen bond contacts formed between M^{pro} residues (white) and the bound TRMT1 peptide (light blue) are illustrated as yellow dashed lines. M^{pro} Phe140, His163, and Glu166 recognize the TRMT1 P1 Gln530 sidechain; additional sidechain and backbone hydrogen bond contacts include M^{pro} Thr24-TRMT1 Thr534, M^{pro} Thr26-TRMT1 Asn532, M^{pro} Asn142-TRMT1 Asn532, M^{pro} Glu166-TRMT1 Arg528, and M^{pro} Gln189-TRMT1 Leu529, consistent with canonical M^{pro}-peptide substrate contacts in the active site. The TRMT1 Gln530-Ala531 peptide bond is positioned between catalytic dyad residues His41 and Cys145Ala (maroon).

A

Substrate	P4	P3	P2	P1	P1′	P2′	P3′	P4′	P2′ Ψ (degrees)
4/5	A	V	L	Q	S	G	F	R	168.1
5/6	V	T	F	Q	S	A	V	K	159.5
6/7	A	T	V	Q	S	K	M	S	103.0
8/9	V	K	L	Q	N	N	E	L	157.3
9/10	V	R	L	Q	A	G	N	A	178.5
10/11	P	M	L	Q	S	A	D	A	157.0
15/16	P	K	L	Q	S	S	Q	A	179.7
TRMT1	**P**	**R**	**L**	**Q**	**A**	**N**	**F**	**T**	**115.5**

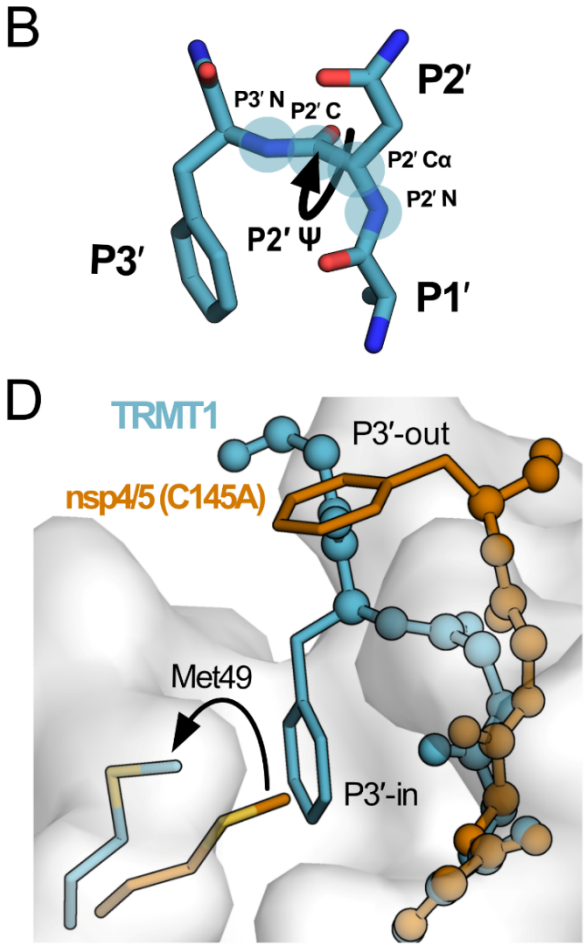

Figure 3. Analysis of M^{pro}-peptide structures illustrates two distinct substrate binding modes. **A)** Comparison of known M^{pro} substrate cleavage sequences and the P2′ Ψ backbone dihedral angles measured in the corresponding C145A M^{pro}-peptide structures for each substrate. We included all known C145A M^{pro}-viral peptide structures in this analysis, except those that were missing the P3′ residue or had poorly-defined electron density for the C-terminal portion of the peptide; structures used in this analysis are PDB IDs: 7MGS, 7T8M, 7DVW, 7T9Y, 7TA4, 7TA7, 7TC4, and 8D35. Additionally, since a C145A M^{pro}-nsp6/7 structure was not available, we used an H41A M^{pro}-nsp6/7 structure (PDB 7VDX) for this analysis. **B)** Section of an M^{pro}-bound peptide substrate showing residues P1′, P2′, and P3′, with the key P2′ Ψ dihedral angle illustrated with a curved arrow; the four backbone atoms that define the P2′ Ψ dihedral angle are labeled and highlighted with blue circles (P2′N–P2′Cα–P2′C–P3′N). **C)** Alignment of peptide substrate backbones in the M^{pro} active site reveals two distinct binding modes at the C-terminal end of the bound peptides characterized by P2′ Ψ dihedral angles ≥ 157° (nsp4/5, nsp5/6, nsp8/9, nsp9/10, nsp10/11, nsp15/16) or ≤ 116° (TRMT1, nsp6/7). Peptide overlays were generated by aligning SARS-CoV-2 M^{pro}-peptide substrate structures in PyMOL. The location of the P2′ Ψ dihedral angle in the substrate peptide backbone is denoted with a star. **D)** Alignment of nsp4/5- and TRMT1-bound M^{pro} structures showing divergent C-terminal peptide substrate binding modes in the M^{pro} active site. The backbone geometry of nsp4/5 (P2′ Ψ = 168°) positions the P3′ Phe sidechain away from the M^{pro} surface ('P3′-out' conformation), while the TRMT1 backbone geometry (P2′ Ψ = 115°) positions the P3′ Phe sidechain toward the M^{pro} active ('P3′-in' conformation) site where it displaces M^{pro} Met49 to open and occupy the S3′ pocket.

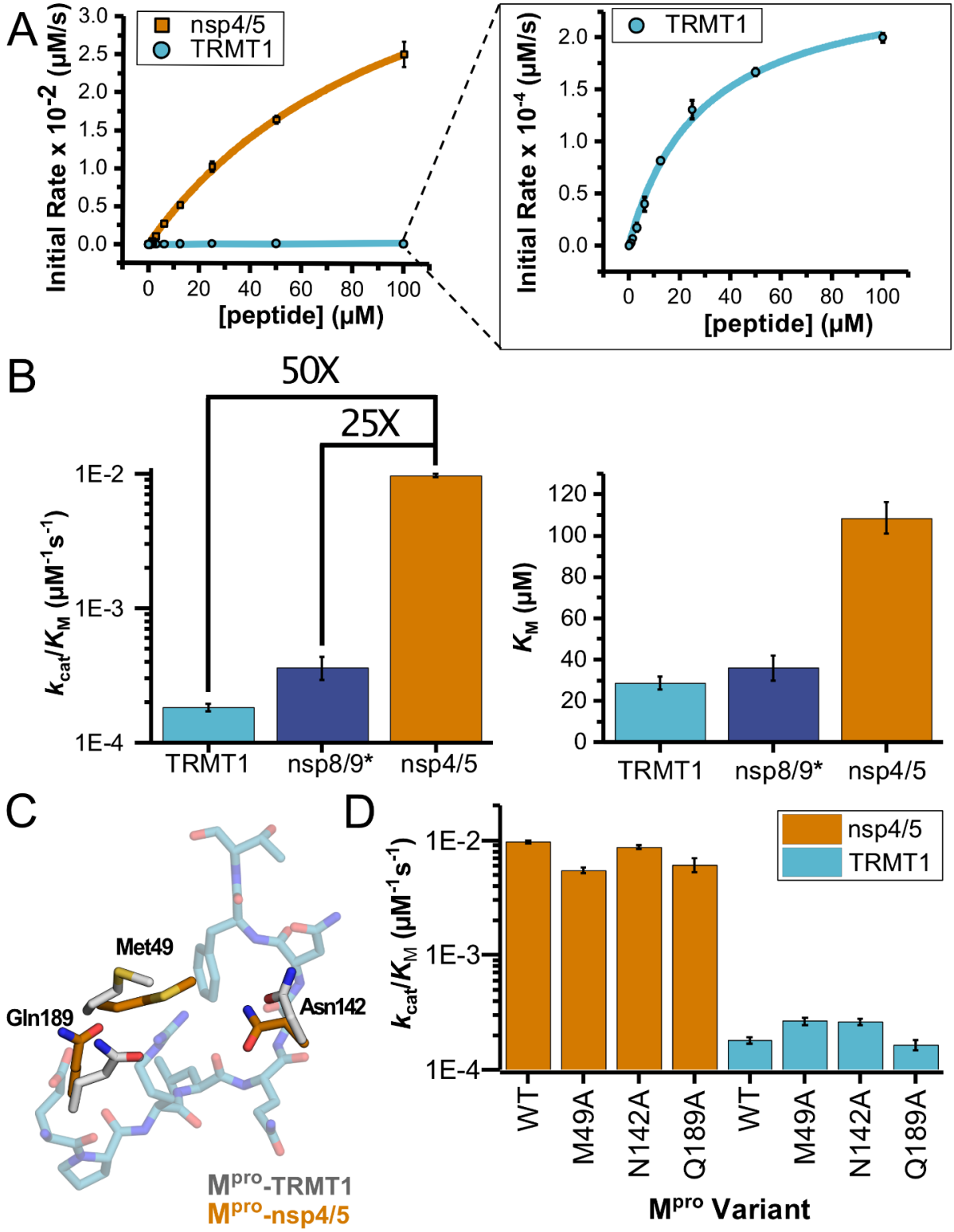

Figure 4. Human TRMT1 peptides are cleaved with similar catalytic efficiencies to known M^{pro} substrates. **A)** Kinetics of nsp4/5 and TRMT1 peptide cleavage by M^{pro}. To initiate the reaction, 50nM enzyme was added to 100-0.097 μM peptide. Each fluorogenic peptide was conjugated with a quenching moiety, and upon peptide cleavage, the fluorescence of the cleavage product was

measured to determine initial rates of the reaction. nsp4/5 rates were faster than those observed with the TRMT1 peptide. **B)** The catalytic efficiency of TRMT1 is similar to that reported of nsp8/9, though both of these substrates exhibit a large difference from nsp4/5. This suggests that TRMT1 is a feasible substrate of M^{pro}. *****nsp8/9 kinetic data are from MacDonald *et al.* (15); these data were measured under similar assay conditions to our nsp4/5 and TRMT1 data and our nsp4/5 kinetic parameters agree closely with those measured by MacDonald *et al*.. **C)** Illustration of changes in M^{pro} Met49, Asn142, and Gln189 residue positioning in TRMT1-bound (white) versus nsp4/5-bound (orange) structures. The TRMT1 peptide is shown in blue; nsp4/5 peptide is not shown. **D)** No major changes in catalytic efficiency are observed for nsp4/5 and TRMT1 peptide cleavage upon mutagenesis of key M^{pro} residues involved in TRMT1 binding and recognition.

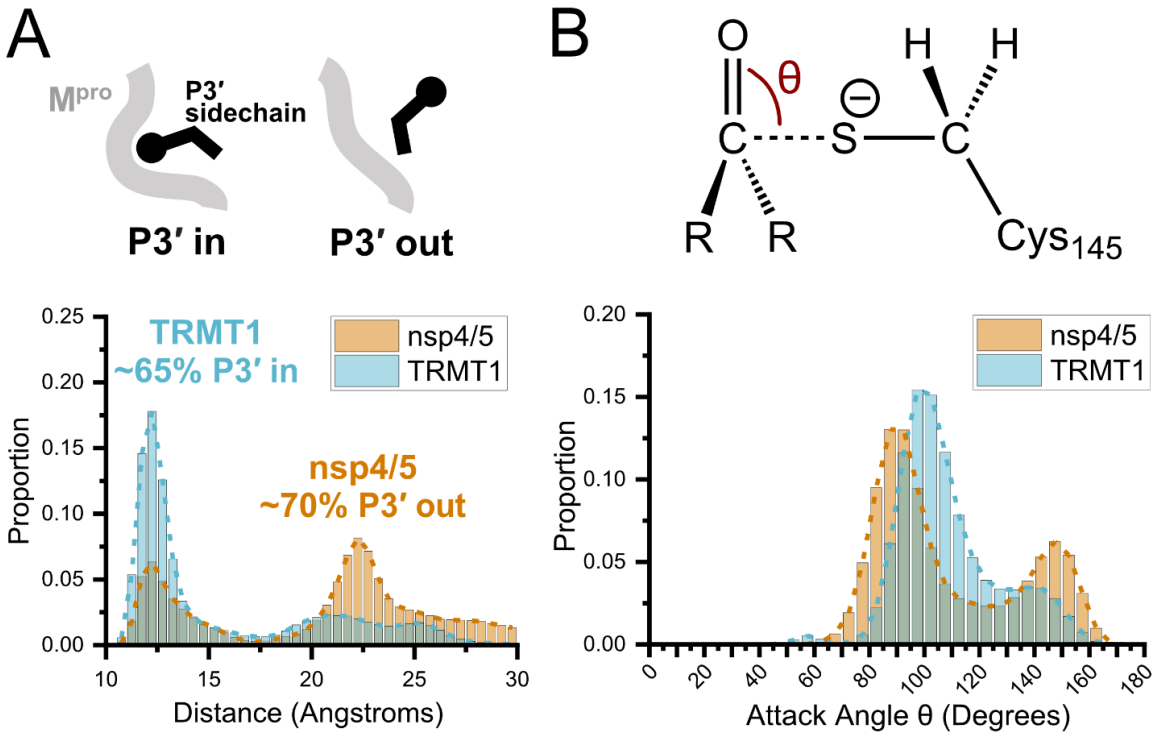

Figure 5. Molecular dynamics (MD) simulations confirm dominant peptide binding conformations and suggest discrimination in cleavage kinetics result catalytic steps that follow initial binding and nucleophilic attack. **A)** Distribution of the sum of the minimum distance for P3′ Phe residue in nsp4/5 or TRMT1 from three residues (Thr25, Met49, Cys44) which form the S3′ subsite; P3′-in and P3′-out conformations are illustrated above the distribution plot. The much larger proportion of TRMT1 at smaller distances reflects the peptide's preference for binding in the P3′-in conformation where TRMT1 P3′ Phe occupies the S3′ pocket during the majority of the MD simulation. **B)** Distribution of the attack angle of the nucleophilic M^{pro} Cys145 sulfur atom and the substrate carbonyl carbon atom in the to-be-cleaved amide bond (S–C=O angle θ, top illustration) during the course of the MD simulation. Although nsp4/5 has a higher proportion of attack angles observed closer to the optimal 90 degrees compared to TRMT1, consistent with faster nsp4/5 cleavage kinetics, this small preference is insufficient to explain the 200-fold faster cleavage kinetics of nsp4/5 observed in experimental proteolysis assays.

Figure 6. The M^{pro}-targeted TRMT1 cleavage site sequence (human TRMT1 residues 526-536) is highly conserved in primates and most mammals, with the notable exception of rodents, where the glutamine Q530 residue most critical for M^{pro}-directed cleavage is substituted to a lysine in Muroidea. Sequence logo plots of the cleavage site in TRMT1(526-536), produced with WebLogo3. The human reference sequence is in black and orange residues show the differences. **A**, **B**, and **C** panels are from primates, mammals, and rodents, respectively.