

Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase

Yu Chen^a, Hui Cai^a, Ji'an Pan^a, Nian Xiang^a, Po Tien^a, Tero Ahola^b, and Deyin Guo^{a,1}

^aState Key Laboratory of Virology and Modern Virology Research Center, College of Life Sciences, Wuhan University, Wuhan 430072, Peoples Republic of China; and ^bInstitute of Biotechnology, University of Helsinki, 00014, Helsinki, Finland

Edited by Paul Ahlquist, University of Wisconsin, Madison, WI, and approved December 30, 2008 (received for review September 4, 2008)

The N7-methylguanosine (m7G) cap is the defining structural feature of eukaryotic mRNAs. Most eukaryotic viruses that replicate in the cytoplasm, including coronaviruses, have evolved strategies to cap their RNAs. In this report, we used a yeast genetic system to functionally screen for the cap-forming enzymes encoded by severe acute respiratory syndrome (SARS) coronavirus and identified the nonstructural protein (nsp) 14 of SARS coronavirus as a (guanine-N7)-methyltransferase (N7-MTase) *in vivo* in yeast cells and *in vitro* using purified enzymes and RNA substrates. Interestingly, coronavirus nsp14 was previously characterized as a 3'-to-5' exoribonuclease, and by mutational analysis, we mapped the N7-MTase domain to the carboxy-terminal part of nsp14 that shows features conserved with cellular N7-MTase in structure-based sequence alignment. The exoribonuclease active site was dispensable but the exoribonuclease domain was required for N7-MTase activity. Such combination of the 2 functional domains in coronavirus nsp14 suggests that it may represent a novel form of RNA-processing enzymes. Mutational analysis in a replicon system showed that the N7-MTase activity was important for SARS virus replication/transcription and can thus be used as an attractive drug target to develop antivirals for control of coronaviruses including the deadly SARS virus. Furthermore, the observation that the N7-MTase of RNA life could function in lieu of that in DNA life provides interesting evolutionary insight and practical possibilities in antiviral drug screening.

RNA capping | exoribonuclease | yeast | alphavirus | flavivirus

Eukaryotic mRNAs possess a 5'-terminal cap structure, in which an N7-methyl-guanine moiety is linked to the first transcribed nucleotide by a 5'-5' triphosphate bridge. Cap formation requires 3 sequential reactions catalyzed by RNA triphosphatase (TPase), guanylyltransferase (GTase), and (guanine-N7)-methyltransferase (N7-MTase) (1). Most cytoplasmically replicating eukaryotic viruses that do not have access to the cellular capping apparatus located in the nucleus have evolved strategies to cap their RNAs. Although the final cap structures of viral and cellular mRNAs are similar, some RNA viruses adopt unique biochemical mechanisms to build the cap structure (2–4).

Coronaviruses are the etiological agents of respiratory and enteric diseases in humans and livestock, exemplified by the life-threatening severe acute respiratory syndrome (SARS) caused by SARS coronavirus (SARS-CoV) (5). The large genomic RNA of coronaviruses contains 2 large ORFs, 1a and 1b, encoding for viral replicase/transcriptase at the 5' terminus and a varied number of ORFs encoding structural and virus-specific accessory proteins at the 3' terminus. The large polypeptides translated from ORFs 1a and 1b are processed into 16 nonstructural proteins (nsps) by 2–3 viral proteinases (6). Strikingly, the coronavirus genome is predicted to encode several RNA processing enzymes that are not present in small RNA viruses, including nsp14 exoribonuclease (ExoN) (7, 8) and nsp15 endoribonuclease (9).

Although the RNAs of coronaviruses and related toroviruses appear to carry a 5'-cap structure (10, 11), the enzymes involved in RNA capping are poorly characterized. Coronavirus nsp13 has been shown to exhibit an RNA TPase activity *in vitro* (12) but a direct role of nsp13 in RNA capping still awaits experimental evidence. Coronavirus nsp16 was predicted to be an S-adenosyl-L-methionine (SAM)-dependent RNA 2'-O MTase (6, 13), and, recently, this activity was confirmed for nsp16 of *Feline coronavirus* (FCoV) (14). However, the N7-MTase and GTase essential for RNA cap formation are currently unknown for coronaviruses. Although a bioinformatic analysis predicted that the SARS-CoV unique domain (SUD) within nsp3 (nsp3-SUD) may possess N7-MTase activity (15), this assumption is weakened by the fact that most coronaviruses do not encode this domain (6). Therefore, identification and characterization of N7-MTase and GTase would be of paramount importance to elucidation of the RNA capping mechanisms of coronaviruses.

Here we used a yeast genetic system to screen for the 3 essential capping enzymes (TPase, GTase, and N7-MTase) possibly encoded by coronaviruses. In previous studies, it has been shown that the capping functions of yeast cells can be replaced by the cap-forming enzymes of mammals and DNA viruses (16–18). In this study, we showed that coronavirus nsp14 could replace the yeast cap MTase *in vivo*, thus providing the first experimental evidence that coronavirus nsp14 acted as an N7-MTase involved in RNA cap formation and demonstrating that a cap-forming enzyme encoded by RNA life could complement the corresponding function of a cellular capping apparatus in DNA life.

Results

Functional Screening for Coronavirus Cap-Forming Enzymes. To functionally screen for coronavirus RNA capping enzymes in yeast cells, we generated a yeast expression plasmid pMceK294A, in which the coding sequence of viral proteins was fused with the N terminus of the GTase domain of the mammalian capping enzyme (amino acid 211–597 of Mce1), so that the fusion protein can be targeted to the RNA polymerase II complex. The GTase activity of Mce1 itself was demolished by a point mutation (K294A). We cloned the coding sequences of all of the nonstructural proteins (nsp1–16) or their individual domains of SARS-CoV strain WHU (19) into pMceK294A and transformed the expression constructs into 3 yeast strains, YBS2, YBS20, and

Author contributions: Y.C., T.A., and D.G. designed research; Y.C., H.C., J.P., N.X., and D.G. performed research; T.A. contributed new reagents/analytic tools; Y.C., H.C., P.T., and D.G. analyzed data; and D.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: dguo@whu.edu.cn.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808790106/DCSupplemental.

© 2009 by The National Academy of Sciences of the USA

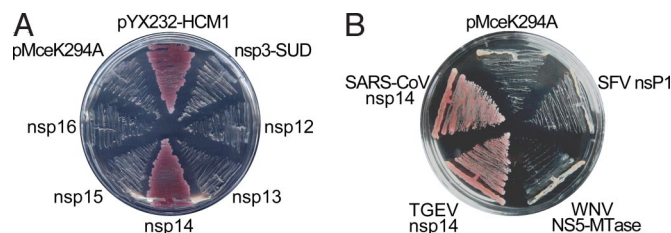


Fig. 1. Functional screening for RNA cap-forming functions encoded by SARS-CoV. (A) YBS40 could be complemented by SARS-CoV-nsp14 and the positive control pYX232-HCM1 but not by other nonstructural proteins of SARS-CoV. (B) The nsp14 of both SARS-CoV and TGEV could rescue the growth of YBS40 on 5-FOA medium. The known N7-MTases including nsP1 of SFV and NS5-MTase domain of WNV could not substitute for the yeast N7-MTase activity.

YBS40, in which the chromosomal genes encoding the essential cap-forming enzymes GTase (*ceg1*), TPase (*cet1*), and N7-MTase (*abd1*) are deleted and cell growth is thus contingent on the maintenance of the corresponding yeast gene on a plasmid carrying *URA3* gene as a selection marker (17, 18). These yeast cells are unable to form colonies on agar medium containing 5-fluoro-orotic acid (5-FOA), which selects against *URA3* plasmid, unless the cells are transformed with a second plasmid bearing a functional homolog from another source. Therefore, the capacity of viral proteins to replace the yeast capping enzymes can be tested by plasmid shuffle and positive growth selection of 5-FOA-resistant colonies.

The system was first validated by introducing mammalian capping enzymes and empty vector pMceK294A into cells of yeast strains YBS20 (*cet1Δ*), YBS2 (*ceg1Δ*), and YBS40 (*abd1Δ*). The mouse capping enzyme Mce1 that possesses both TPase and GTase activities could rescue the cell growth of YBS20 and YBS2 while the human N7-MTase Hcm1 could complement the growth of YBS40 on 5-FOA-containing medium (Fig. S1). In contrast, the expression vector pMceK294A bearing a mutated GTase domain of Mce1 could not substitute for any of the yeast capping functions. These data indicated that the screening system for capping enzymes was reliably established.

In total, 18 expression constructs of SARS-CoV nonstructural proteins were generated and transformed into the 3 yeast strains. The resulting yeast transformants were streaked on 5-FOA-containing medium for plasmid shuffle and functional screening at 20 °C, 25 °C, and 30 °C, respectively. After extensive screening and testing, we found that SARS-CoV nsp14 could complement the cell growth of YBS40 (Fig. 1A), but there was no obvious growth complementation of YBS2 and YBS20 by any of the SARS-CoV proteins (data not shown). The same results were obtained when the yeast cells were incubated at different temperatures.

Coronavirus nsp14 Is Able to Complement the Deletion of Cellular Cap N7-MTase. As shown in Fig. 1A, the SARS-CoV nsp14 was able to replace the yeast cap MTase to sustain cell growth as effectively as human N7-MTase (Hcm1), but all other SARS-CoV proteins including nsp12, nsp13, nsp15, nsp16, and nsp3-SUD were not. The negative data with other SARS-CoV proteins are not shown.

We further tested whether the nsp14 of a different coronavirus and known cap MTases of other RNA viruses could be functional in lieu of Abd1 of yeast cells. As shown in Fig. 1B, the nsp14 of transmissible gastroenteritis coronavirus (TGEV), a member of group 1 coronaviruses that are distantly related to SARS-CoV, was able to complement the cell growth of YBS40 as efficiently as SARS-CoV nsp14 did, suggesting that MTase function of nsp14 might be conserved among coronaviruses.

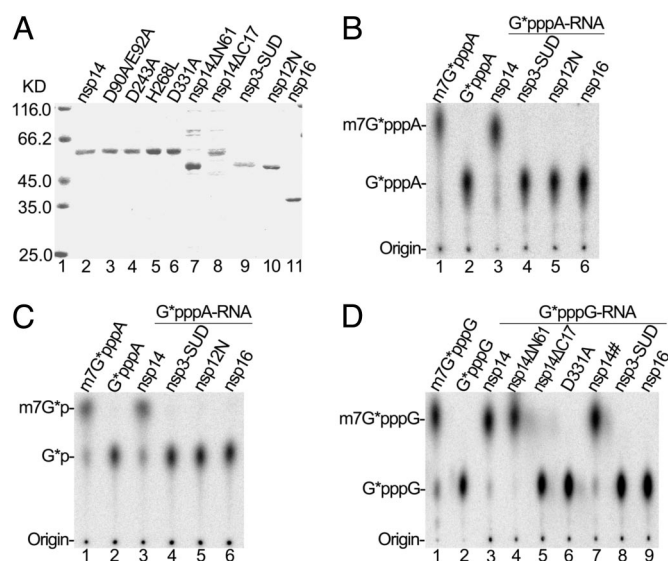


Fig. 2. Biochemical analyses of the guanine-N7 methyltransferase activity of nsp14 of SARS-CoV. (A) Purification and SDS/PAGE analysis of recombinant nsp14 (lane 2), point mutants (lanes 3–6), truncation mutants of nsp14 (lanes 7–8), nsp3-SUD domain (lane 9), nsp12N (lane 10), and nsp16 (lane 11) of SARS-CoV. nsp12N represents a distinct N-terminal domain without predicted function in nsp12. The sizes of protein markers (lane 1) are indicated on the Left. (B) TLC analysis of nuclease P1-resistant cap structures released from the G*pppA-RNA treated by nsp14, nsp3-SUD, nsp12N, and nsp16, respectively (lanes 3–6). The asterisk indicates that the following phosphate is 32 P-labeled. The positions of origin and migration of m7G*pppA/G*pppA (lanes 1–2) are indicated on the Left. The positive controls for m7G*pppA and G*pppA were generated by vaccinia capping enzyme. (C) TLC analysis of tobacco acid pyrophosphatase-digested products from the reactions in panel B. The positions of origin, m7G*p and G*p are indicated on the Left. (D) TLC analysis of nuclease P1-resistant cap structures released from the G*pppG-RNA treated by nsp14, nsp14ΔN61, nsp14ΔC17, mutant D331A, nsp14# (stored in elution buffer, otherwise identical to nsp14 in lane 3), nsp3-SUD, nsp16, respectively (lanes 3–9). The positions of origin and migration of m7G*pppG/G*pppG (lanes 1–2) are indicated on the Left.

Semliki Forest virus (SFV) nsP1 (4) and West Nile virus (WNV) NS5 (20) are well-characterized viral MTases, but in fusion with mouse GTase domain, they could not replace the cap MTase activity in yeast (Fig. 1B). These observations may reflect the fact that SFV adopts an unconventional capping mechanism (4) while WNV MTase is viral sequence specific (21). Taken together, these data suggest that coronavirus nsp14 may function as cap MTase in a sequence-nonspecific manner, which is distinct from that of other positive-stranded RNA viruses.

Biochemical Evidence that Coronavirus nsp14 Acts as Cap N7-MTase.

To confirm the cap MTase activities of coronavirus nsp14, recombinant SARS-CoV nsp14 and various mutants with an N-terminal 6xHis tag were expressed in *Escherichia coli* cells and purified (Fig. 2A). The substrate RNAs contained an unmethylated G*pppN cap (where the * indicates that the following phosphate was 32 P-labeled). One substrate RNA with a GpppA cap represented the 5'-terminal 259 nucleotides of SARS-CoV genome, and the other was an unrelated GpppG-capped RNA of 52 nt. After treatment of the substrate RNA with viral proteins, the RNA was digested with nuclease P1 (which cleaves capped RNAs into 3'-OH-terminated cap structures and 5'-pN_{OH}), and the resulting N7-methylated cap structure (m7G*pppA) could be readily separated from nonmethylated cap structure (G*pppA) by TLC as previously shown in flavivirus MTase assays (20).

As shown in Fig. 2B, treatment of G*pppA-capped RNAs with SARS-CoV nsp14 (Fig. 2A, lane 2) in the presence of SAM

	Core ExoN Domain	Core MTase Domain	YBS40 Growth	MTase Activity	ExoN Activity
wt	aa: 1 90 92 243 268 331 333 527	DxE D H DxG	++	++	++
D90A/E92A	1 AxA 527		++	++	—
D243A	1 A 527		++	++	—
H268L	1 L 527		++	++	+
D331A	1 AxG 527		—	—	++
ΔN31	32 527		++	N.D.	N.D.
ΔN61	62 527		++	++	N.D.
ΔN78	79 527		+	N.D.	N.D.
ΔN90	91 527		—	N.D.	N.D.
ΔC6	1 521		+	N.D.	N.D.
ΔC17	1 510		—	—	N.D.
ΔC27	1 500		—	N.D.	N.D.

Fig. 3. The domain structure of SARS-CoV nsp14 and its mutants. The core domains of ExoN and MTase are indicated with open box and solid black box, respectively. The conserved residues DxG (amino acid 331–333) in the MTase domain are putative SAM-binding sites. The conserved residues D-E-D-H in the ExoN domain are indicated. The truncation mutants are named as the number of the deleted amino acids from either N or C terminus. The capability for complementing yeast growth and the activities of MTase and ExoN are presented in the *Right* panel. The values corresponding to ++, + and – are >60% activity of wild-type protein (or wild-type growth), >15% (or slow growth), and <15% (or no growth). ND, not determined.

resulted in P1 cleavage products (Fig. 2*B*, lane 3) that comigrated with m⁷G*pppA (Fig. 2*B*, lane 1). As a positive control, the m⁷G*pppA was produced by the commercially available vaccinia D1/D12 capping enzyme. In contrast, treatment of substrate RNA with nsp3-SUD, nsp12N, and nsp16 did not generate a similar product corresponding to the position of m⁷G*pppA on the TLC plate (Fig. 2*B*, lanes 4–6). These results demonstrated that the nsp14 could methylate the GpppA cap of RNAs, consistent with the results of *in vivo* assays in yeast cells.

The capped RNA substrates and products described above (Fig. 2*B*) were also digested with tobacco acid pyrophosphatase (TAP) that releases m⁷G*p and G*p from m⁷G*pppA and G*pppA-capped RNAs, respectively. As expected, TAP digestions of RNA products treated with nsp14 resulted in a ³²P-labeled band on the TLC plate (Fig. 2*C*, lane 3) that corresponded to m⁷G*p, a positive control generated by vaccinia capping enzyme (Fig. 2*C*, lane 1). In contrast, TAP digestions of G*pppA-capped RNA treated with other proteins did not generate any signal at the position of m⁷G*p (Fig. 2*C*, lanes 4–6). These data further confirmed that SARS-CoV nsp14 could methylate GpppA-capped RNA at the N7 position of the capping guanosine.

The N7-MTase activity of nsp14 appeared to be sequence-nonspecific *in vivo* as it was able to sustain the growth of yeast cells. Therefore, we further tested in biochemical assays whether SARS-CoV nsp14 was able to methylate the cap structure of a different substrate RNA that possessed a GpppG cap. The G*pppG-RNA was treated with nsp14 and control proteins in a manner similar to G*pppA-RNA. As shown in Fig. 2*D*, nsp14 could modify the GpppG cap structure in the presence of SAM, resulting in the P1 digestion product m⁷G*pppG (Fig. 2*D*, lane 3) similarly to the positive control (Fig. 2*D*, lane 1). These results indicated that coronavirus nsp14 was a sequence-nonspecific cap MTase.

The N7-MTase Core Domain Is Located at the C Terminus and Coupled with an ExoN Domain in nsp14. Because nsp14 has been shown to possess exoribonuclease activity, we wanted to address the domain location of N7-MTase. The ExoN domain was predicted to reside in the N-terminal part of nsp14 (6), and we predicted that the MTase domain was located in the C-terminal half of nsp14 (Fig. 3, *Left* panel). By structure-based sequence alignment, conserved secondary structure elements and motif I (DxG) implicated in

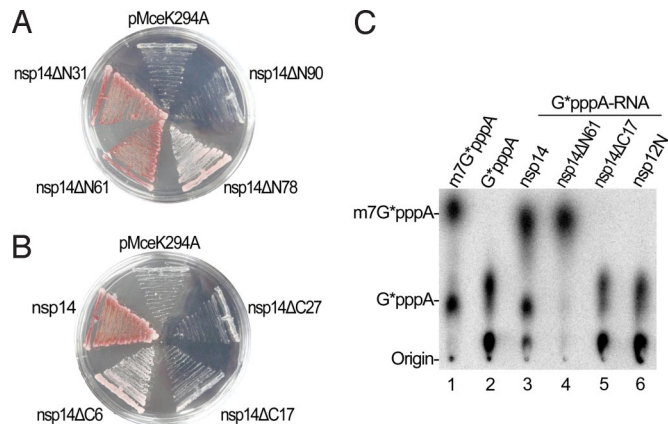


Fig. 4. MTase activity analysis of truncation mutants of SARS-CoV nsp14. (A) YBS40 (*ade2*) transformed with the N-terminal truncation mutants of SARS-CoV nsp14. ΔN31 and N61 could readily rescue the cell growth of YBS40 on 5-FOA medium while cells rescued by ΔN78 grew slower than that with ΔN31 and ΔN61 as shown by lesser red pigment accumulation and smaller cell density. In contrast, ΔN90 and other N-terminal truncation lost the MTase activity in yeast. (B) YBS40 transformed with the C-terminal truncation mutants of SARS-CoV nsp14. The mutant ΔC6 (amino acid 1–521) could rescue YBS40 on 5-FOA medium but the cells grew obviously slower than those with wild-type nsp14. All other C-terminal truncations abolished the MTase activity of nsp14 in yeast cells. (C) TLC analysis of nuclease P1-resistant cap structures released from the G*pppA-RNA methylated by nsp14, nsp14ΔN61, nsp14ΔC17, nsp12N, respectively (lanes 3–6). The positions of origin and migration of m⁷G*pppA/G*pppA (lanes 1–2) are indicated on the *Left*.

SAM-binding/catalysis of N7-MTase could be recognized in the C-terminal domain of nsp14 (Fig. 3 and Fig. S2).

To delineate the MTase domain in nsp14, we created a series of deletion mutants by truncating from either the N terminus or C terminus of nsp14 (Fig. 3). These mutants were introduced into YBS40 cells for functional assays. As shown in Fig. 4*A*, the mutants ΔN31 and ΔN61 were able to complement yeast growth as efficiently as the wild-type nsp14 while the capability of ΔN78 was significantly attenuated. In contrast, N-terminal mutants with longer truncations (≥90 amino acids) could not complement the growth of yeast YBS40 (Fig. 4*A*).

At the C-terminal end, deletion of 6 amino acids (ΔC6) already severely impaired the MTase activity in yeast cells whereas the deletions of 17 amino acids (ΔC17) or more (ΔC27) led to complete loss of the capability to complement yeast growth (Figs. 3 and 4*B*). Some of these mutants were also tested in biochemical assays and the results were in accordance with *in vivo* analysis (Figs. 4*C* and 2*D*, lanes 4–5). As the N-terminal deletion mutants that retained the MTase activity still contained the conserved ExoN domain, these data did not define the exact location of MTase domain in nsp14.

We then tested point mutations located in the conserved motifs of the ExoN domain (D90A/E92A, D243A, and H268L) and the predicted MTase domain (D331A) (Fig. 3). The nsp14 mutants bearing these point mutations were first tested in the yeast system (Fig. 5*A*). The results showed that the D331A mutation destroyed the MTase activity of nsp14 *in vivo*; in contrast, all of the point mutations in the ExoN domain did not significantly interfere with the MTase function in yeast cells (Fig. 5*A*). These results were further verified by the *in vitro* biochemical analysis (Fig. 5*B*). In another set of assays, we tested the effects of the point mutations on ExoN activity. The D331A mutation did not influence the ExoN activity whereas the point mutations in the ExoN domain attenuated the exoribonuclease activity to varied degrees (Fig. 5*C* and *D*). The results of the *in vivo* and *in vitro* tests of various truncation and point mutations

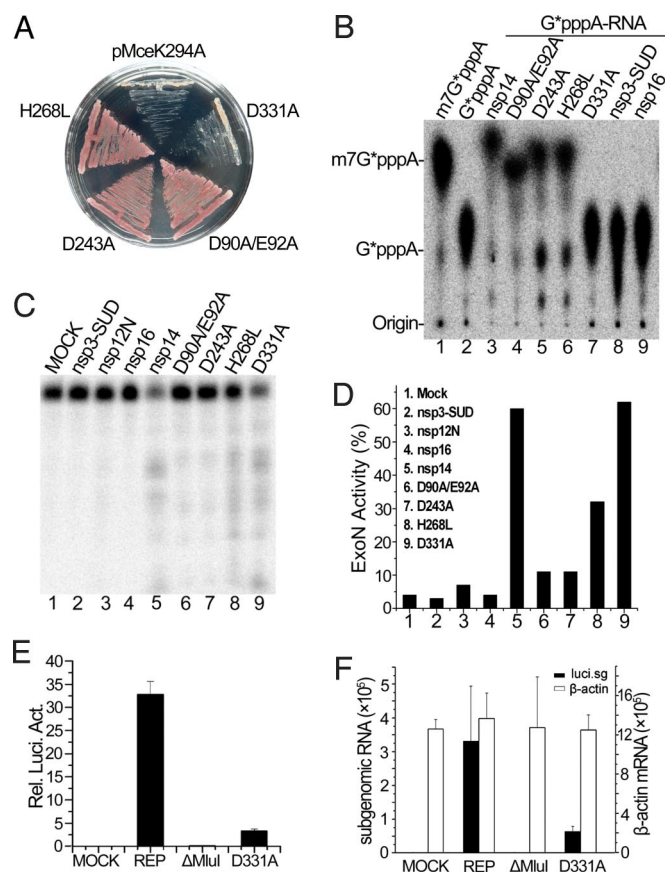


Fig. 5. Effects of MTase and ExoN active-site mutations on enzymatic activities of nsp14 and on SARS-CoV replicon. (A) YBS40 was transformed with the point mutants of SARS-CoV nsp14. D90A/E92A, D243A, and H268L mutations in the ExoN domain could rescue the cell growth of YBS40 on 5-FOA medium while D331A in the predicted MTase domain lost the MTase activity in yeast cells. (B) TLC analysis of nuclease P1-resistant cap structures released from the G*pppA-RNA treated by wild-type nsp14 and its mutants D90A/E92A, D243A, H268L, and D331A (lanes 3–7). The nsp3-SUD and nsp16 (lanes 8–9) were used as negative controls. The positions of origin and migration of m7G*pppA/G*pppA (lanes 1–2) are indicated on the Left. (C) Analysis of 3'-to-5' ExoN activity of nsp14 and its mutants D90A/E92A, D243A, H268L, and D331A (lanes 5–9) in 10% urea-PAGE using single-stranded RNA substrates that were ³²P-labeled at the 5'-end. The mock treatment or with nsp3-SUD, nsp12N, and nsp16 were used as negative controls (lanes 1–4). (D) Quantitation of ExoN activity by analyzing the band intensity shown in Fig. 5C with ImageQuant software. The values in the y axis indicate the ratios of digestion products to input RNA within each lane. All products migrating below full-length substrate were quantitated together as digestion products. Input RNA represents digestion products plus undigested full-length substrate. (E) Effect of D331A mutation on luciferase reporter of SARS-CoV replicon. The reporter-carrying replicon (rep) and a deficient deletion mutant (ΔMlul) were described previously (22). (F) Quantitative PCR of subgenomic RNAs with the same set of samples as Fig. 5E (Left y axis). β-Actin mRNA was detected as internal control to show that equal amount of RNA was used (Right y axis). Graphs show the mean of RNA copy number of 3 experiments ± SD.

are summarized in Fig. 3, and these data show that the core MTase domain is located in the C-terminal half of nsp14.

N7-MTase Is Important for SARS-CoV Genome Replication/Transcription. We further addressed whether N7-MTase is required for replication. The D331A substitution was introduced into a SARS-CoV replicon that carries a luciferase reporter, and decrease in reporter activity indicates that RNA genome replication and/or transcription is impaired (22). The results showed that the D331A mutation led to reduction of luciferase activity

by 90% and the copy number of subgenomic RNA by 81% (Fig. 5E and F). These results indicate that the N7-MTase plays an important role in SARS-CoV replication/transcription.

Discussion

Coronaviruses possess the largest genome known among RNA viruses and employ unusually complicated mechanisms for genome replication and expression. The 16 viral nonstructural proteins are translated directly from the genomic RNA presumably by a cap-dependent mechanism (10). In this work, we identified SARS-CoV nsp14 as an N7-MTase involved in RNA cap formation by functional screening in yeast cells and by biochemical assays. However, this finding does not rule out the possibility that SARS-CoV encodes another N7-MTase, such as nsp3-SUD (15). Although the nsp14 N7-MTase acted in a sequence-unspecific manner in vitro, it may only have access to viral RNAs as viral replication takes place in association with double membrane vesicles (23).

Identification of coronavirus nsp14 as cap MTase was unexpected in the background that nsp14 had been characterized as an exoribonuclease (7, 8) and that 2 proteins other than nsp14 had been predicted to contain a conserved domain for MTase (6, 13–15). In this study, the nsp14 of 2 coronaviruses belonging to distantly related groups in the *Coronaviridae* was shown to be able to functionally replace yeast cap MTase, but the N7-MTases of SFV and WNV in fusion with mouse GTase domain were not. These observations probably reflect the diversified nature of the organization and mechanisms of RNA virus-capping apparatus. The SFV capping enzyme nsP1 possesses both MTase and GTase activities (24) and N7 methylates free GTP before it is transferred by GTase to 5'-end of the viral RNA (4). The WNV MTase is physically linked with the viral RNA-dependent RNA polymerase in NS5 and acts both as N7 and 2'-O MTase in a sequence-specific manner (25). However, it is also possible that the SFV and WNV fusion proteins do not fold correctly to an active conformation or do not attain their proper localization in yeast nucleus. The coronavirus N7-MTase identified in this study exhibits a completely different organization from that of any other described MTase as it is physically linked with an ExoN domain that, among RNA viruses, only exists in coronaviruses and related toroviruses and roniviruses. These observations suggest that coronavirus nsp14 may represent a unique form of MTase that is distinct from that of any other virus groups. Nevertheless, coronavirus nsp14 is functionally equivalent to yeast cap MTase, and this is the first evidence that a capping function of cytoplasmic RNA life can replace that of DNA life in the nucleus. This might reflect the converged evolution of capping functions of RNA viruses and eukaryotic cells or the evolutionary transfer of some of the capping functions across domain boundaries.

The RNAs of many eukaryotic viruses bear a cap-1 structure that results from methylation at the 2'-O position of the first transcribed nucleotide in the cap-0 structure by 2'-O MTase (1). Very recently, the 2'-O MTase activity of FCoV was mapped to nsp16, suggesting that coronaviruses may adopt a cap-1 structure in their mRNAs (14). Our observation that coronavirus nsp14 was functional in lieu of cellular N7-MTase, together with the recent finding that FCoV 2'-O MTase prefers cap-0 structure as substrate for generating cap-1 structure (14), strongly suggests that the cap methylation in coronaviruses may follow the canonical order of capping steps found for eukaryotic mRNAs, although the modular structure and composition of the capping apparatus is clearly different.

SAM-dependent MTases belong to a large family of enzymes with the Rossmann-like fold, but the profound conservation is not reflected in the primary sequence (26, 27). The core SAM-MTase fold incorporates alternating β-strands (β1–β7) and α-helices (αZ and αA–αE), which form a 7-stranded β-sheet

with 3 helices on each side. Structure-based alignments suggested that the C-terminal half of SARS-CoV nsp14 shares this general fold (Figs. S2 and S3).

By superimposition of the predicted nsp14 MTase structure with the crystal structure of cellular N7-MTase, structural conservation between nsp14 MTase and cellular N7-MTase could be predicted (Fig. S3). The most conserved sequence block of MTases is located in the SAM-binding site that has a sequence signature of D/ExGxGxG (where x means any amino acid) in motif I (26, 27). By sequence and structural alignments, the sequence motif DxGxPxG/A in coronavirus nsp14 could be recognized as the motif corresponding to the D/ExGxGxG motif in N7-MTase (Fig. S2), which thus may represent the SAM-binding site in nsp14 MTase.

Point mutation analysis of the conserved residues in the ExoN and MTase domain of nsp14 revealed that the active sites of the 2 enzymes are functionally distinct and physically independent. However, deletion mutations revealed that the ExoN domain is also important for the MTase activity of nsp14, suggesting that the 2 functional domains may be structurally linked (e.g., by interdigitation of some regions of the domains). Alternatively, a shared functional region, such as an RNA-binding domain, could be proposed for MTase and ExoN. Crystallization and resolution of the atomic structure are needed to finally elucidate the structure-function relationships of coronavirus nsp14.

Some roles for nsp14 in the coronavirus life cycle have been described previously. Deletion or mutation of the conserved residues in the ExoN domain has a lethal or attenuating effect on viral replication (8, 28) and results in reduced fidelity of replication in MHV (29). These mutations abolished or reduced the ExoN activity but did not influence the MTase activity as shown in this report. Point mutations that led to temperature-sensitive phenotype (30) and attenuation in pathogenesis of MHV (31) have been mapped to the C-terminal domain of nsp14 although it is not known whether these mutations are correlated with a change of MTase activity. In this study, we provided direct evidence that nsp14 MTase activity is important for coronavirus replication and/or transcription by showing that the D331A mutation, which abolished the N7-MTase activity, severely impaired the replication/transcription of SARS-CoV replicon.

In summary, our current work identified coronavirus nsp14 as a cap MTase that was functional in yeast cells. The unique structure and configuration of coronavirus nsp14 with both MTase and ExoN activities also provides an attractive target for drug design, and the availability of isogenic yeast strains containing SARS-CoV nsp14 versus human cap MTase will facilitate antiviral drug screening aimed at blocking the cap formation of coronaviruses and effective control of newly emerging and reemerging coronaviruses such as the life-threatening SARS-CoV.

Materials and Methods

Construction of Yeast Expression Plasmids. The sequence encoding the mouse GTase domain (amino acid 211–597) was PCR amplified from pYX1-MCE1 (16). To remove the GTase activity, the K294A mutation was introduced into the GTase sequence by PCR amplification with mutagenic primers using overlap extension PCR. The PCR fragment was digested with NheI and NdeI and inserted into pYX1-MCE1 resulting in pYX1-GT. Subsequently, the EcoRI-NdeI fragment was excised from pYX1-GT and inserted into pYX232-HCM1 (16) in lieu of the Hcm1 gene, resulting in pMceK294A (2 μ m, TRP1), in which BamHI and XhoI cloning sites are available for cloning foreign genes that are fused in-frame with an N-terminal leader peptide (MGSHHHHHHSGHMR1) and the C-terminal GTase domain with K294A mutation, and the expression is under the control of the yeast *TP1* promoter.

Nonstructural proteins (nsp1–16) or their individual domains of SARS-CoV, nsp14 of TGEV, nsp1 of SFV, and N55 MTase domain of WNV, were PCR amplified from cDNAs of SARS-CoV strain WHU (19), TGEV, SFV, and WNV and inserted into pMceK294A, respectively. All of the primers used are listed in Table S1. The truncation mutants of nsp14 of SARS-CoV were generated by PCR using nested primers. The D90A/E92A, D243A, H268L, and D331A mutations of nsp14 were

introduced by overlap PCR with mutagenic primers (Table S2). The PCR fragments were cloned into pMceK294A as described for wild-type nsp14. All of the clones and mutations were confirmed by DNA sequencing.

Yeast Strains and Functional Screening. Yeast strain YBS2 lacks the chromosomal *CEG1* locus encoding yeast GTase and its growth depends on maintenance of plasmid pGYCE-360 (*URA3*) (18). Strain YBS20 lacks the chromosomal *CET1* locus encoding yeast RNA TPase and its growth depends on maintenance of p360-CET1 (*URA3*) (18). Strain YBS40 is deleted at the chromosomal *Abd1* locus encoding the yeast cap MTase (16) and its growth depends on the maintenance of plasmid p360-ABD1 (*URA3*).

For screening the cap-forming functions encoded by SARS-CoV, yeast cells were transformed with the expression plasmids carrying SARS-CoV nonstructural proteins. Trp⁺ transformants were selected at 30 °C on agar medium lacking tryptophan and the cells were then streaked on agar medium containing 0.75 mg/mL of 5-FOA to counterselect *URA3* plasmids at 30 °C, 25 °C, and 20 °C, respectively. The plates were incubated for up to 8 days and formation of FOA-resistant colonies indicated the transformed plasmid carried the function that could replace or complement the endogenous capping function residing in the *URA3* plasmid. The SARS-CoV mutant proteins and other viral proteins were tested similarly.

Cloning, Expression, and Purification of Recombinant Proteins. The coding sequences for nsp14, N-terminal domain of nsp12 (nsp12N), nsp3-SUD, and nsp16 of SARS-CoV were PCR amplified from cDNA sequence of SARS-CoV strain WHU (19) using the primers as listed in Table S3, and inserted into plasmid pET30(a) (Novagen) or pHAT (24). The truncation mutants and D90A/E92A, D243A, H268L, and D331A mutations of nsp14 were generated as described for yeast expression plasmids. All constructs were verified by DNA sequencing and their expression gave rise to recombinant proteins with an N-terminal 6xHis tag. The protein expression and purification were described previously (7).

Preparation of Capped RNA Substrates. RNA substrates representing the 5'-terminal 259 nucleotides of the SARS-CoV genome were in vitro transcribed from PCR products generated using the primers 5'-CAGTAATACGACTCATTATAGATTAGGTTTTTACCTACCCAG-3' and 5'-CTTCGGTCACACCCGGAC. A bacteriophage T7 class II ϕ 2.5 promoter (in italics) was inserted into the forward primer to synthesize the ATP-initiated RNA (underlined) using T7 RNA polymerase (32). To increase the transcription efficiency, the second nucleotide that is U in the SARS-CoV genome was mutated to G in the substrate RNA. A nonviral RNA substrate comprising 52 nucleotides (with G as the first nucleotide) was in vitro transcribed from linear pGEM-T vector (Promega) using T7 RNA polymerase. RNAs containing ³²P-labeled cap structures (m7G*pppA or m7G*pppG and G*pppA or G*pppG) were prepared using a vaccinia virus capping enzyme following the manufacturer's protocol (Epicentre), except that the methyl donor SAM was absent for preparing the G*pppA/G*pppG-RNA, and inorganic pyrophosphatase (0.05 units) (New England Biolabs) was added to enhance the yield of ³²P-labeled G*pppA/G*pppG-RNA (20) because the capping reaction is reversible in the absence of the methyl donor SAM. The m7G*pppA/m7G*pppG-RNA or G*pppA/G*pppG-RNA was then purified twice through Sephadex G-50 spin columns (Roche), extracted with phenol-chloroform, and precipitated with ethanol.

Biochemical Assays for MTase Activity. Purified recombinant proteins (0.5–1 μ g) and 1×10^3 cpm of ³²P-labeled G*pppA/G*pppG-RNA substrates were added to 7.5 μ L reaction mixture [50 mM Tris-HCl (pH 8.0), 6 mM KCl, 2 mM DTT, 1.25 mM MgCl₂, 10 units RNase inhibitor, 0.2 mM SAM] and incubated at 37 °C for 1.5 h. RNA cap structures were liberated with 5 μ g of nuclease P1 (Sigma) in 10 mM Tris-HCl, pH 7.5, 1 mM ZnCl₂ at 50 °C for 30 min, then spotted onto polyethyleneimine cellulose-F plates (Merck) for TLC, and developed in 0.4 M ammonium sulfate. The extent of ³²P-labeled cap was determined by scanning the chromatogram with a PhosphorImager. For tobacco acid pyrophosphatase analysis, the P1 digestion products were extracted with phenol-chloroform and precipitated with ethanol, and subsequently digested by TAP by following the protocols recommended by the manufacturer (Epicentre).

Exoribonuclease Assays. Purified recombinant proteins (0.5–1 μ g) and 300 cpm of the ³²P-labeled single-strand RNA of 52 nucleotides were added to 10 μ L reaction mixture [50 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 50 mM KCl, 50 mM NaCl, 2 mM DTT, 0.1 mg/ml BSA, 10 units RNase inhibitor]. The mixture was incubated at 37 °C for 1.5 h and analyzed in 10% polyacrylamide gel of 20 \times 20 cm with 7.5 M urea. The amount of ³²P-labeled RNA was determined by scanning the gel with a PhosphorImager.

