doi:10.1016/j.jmb.2009.09.040 J. Mol. Biol. (2009) 394, 544–557

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Coronavirus N Protein N-Terminal Domain (NTD) Specifically Binds the Transcriptional Regulatory Sequence (TRS) and Melts TRS-cTRS RNA Duplexes

Nicholas E. Grossoehme[1]†, Lichun Li[2]†, Sarah C. Keane[1], Pinghua Liu[3], Charles E. Dann III[1], Julian L. Leibowitz[3] and David P. Giedroc[1]*

[1]*Department of Chemistry, Indiana University, Bloomington, IN 47405-7102, USA*

[2]*Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843-2128, USA*

[3]*Department of Microbial and Molecular Pathogenesis, Texas A&M University System College of Medicine, College Station, TX 77843-4467, USA*

Received 1 August 2009; received in revised form 11 September 2009; accepted 16 September 2009
Available online
24 September 2009

Edited by D. E. Draper

All coronaviruses (CoVs), including the causative agent of severe acute respiratory syndrome (SARS), encode a nucleocapsid (N) protein that harbors two independent RNA binding domains of known structure, but poorly characterized RNA binding properties. We show here that the N-terminal domain (NTD) of N protein from mouse hepatitis virus (MHV), a virus most closely related to SARS-CoV, employs aromatic amino acid-nucleobase stacking interactions with a triple adenosine motif to mediate high-affinity binding to single-stranded RNAs containing the transcriptional regulatory sequence (TRS) or its complement (cTRS). Stoichiometric NTD fully unwinds a TRS-cTRS duplex that mimics a transiently formed transcription intermediate in viral subgenomic RNA synthesis. Mutation of the solvent-exposed Y127, positioned on the β-platform surface of our 1.75 Å structure, binds the TRS far less tightly and is severely crippled in its RNA unwinding activity. In contrast, the C-terminal domain (CTD) exhibits no RNA unwinding activity. Viruses harboring Y127A N mutation are strongly selected against and Y127A N does not support an accessory function in MHV replication. We propose that the helix melting activity of the coronavirus N protein NTD plays a critical accessory role in subgenomic RNA synthesis and other processes requiring RNA remodeling.

Keywords: nucleocapsid protein; coronavirus replication; SARS coronavirus; protein-RNA interactions; transcriptional regulatory sequence

Introduction

Coronaviruses (CoVs) harbor very large positive sense RNA genomes (≈30 kb) and cause a range of upper and lower respiratory tract infections in both veterinary animals and humans. CoVs include the causative agent of severe acute respiratory syndrome (SARS), SARS-CoV, which was associated with substantial mortality during the initial outbreak originating in Guangdong province in 2002.[1] Two additional human CoVs, NL63 and HKU1, have since been isolated and are associated with lower respiratory tract disease but limited mortality in healthy individuals.[2,3] Mouse hepatitis virus (MHV), on the other hand, is the prototype group 2 coronavirus that serves as a well-studied model system for SARS-CoV.[4]

SARS-CoV and related viruses are highly recombinogenic due in part to an unusual transcription/replication cycle involving the synthesis of 5-8 subgenomic RNA (sgRNA) intermediates.[5] The minus-sense discontinuous transcription model,[6,7] a well accepted model for sgRNA synthesis, postulates that genome circularization, for which there is now genetic evidence,[8] is a requisite initial step. This enables individual transcriptional regulatory sequences (TRSs) in the genome body (TRS-B) to come in close physical proximity to the single TRS in

---

*Corresponding author. E-mail address: giedroc@indiana.edu.

† Contributed equally to this work.

Abbreviations used: N, nucleocapsid protein; TRS, transcription regulatory sequence; SARS, severe acute respiratory syndrome; MHV, mouse hepatitis virus; IBV, infectious bronchitis virus; TGEV, transmissible gastroenteritis virus, sg, subgenomic.

the leader (TRS-L) of the 5′ untranslated region (UTR) during nascent strand synthesis. At each TRS-B sequence, the replication complex either continues transcription of genomic RNA template, or engages in a "template switch" that generates minus-sense sgRNAs. These can subsequently function as templates for production of positive-sense sgRNA transcripts that are ultimately translated by host ribosomes into the structural proteins of the virus.

Coronaviral N is a multifunctional phosphoprotein[9,10] that plays a primary structural role in packaging the RNA genome into a helical ribonucleoprotein,[11] as well as regulatory roles in viral RNA synthesis (replication and transcription), translation, and modulation of host cell metabolism.[12-15] *In situ* crosslinking and immunoprecipitation experiments reveal that N interacts with multiple regions of positive-and negative-sense coronaviral genome and all sgRNAs[12,14] including the 5′ leader. N-specific antibodies inhibit mouse hepatitis virus (MHV) RNA synthesis *in vitro*[16] and N has been shown to significantly enhance the efficiency of RNA replication.[17] These and other data implicate N as an important accessory factor in discontinuous transcription.[18] These multifunctional properties of CoV N are analogous to HIV-1 nucleocapsid protein[19] and make N an attractive antiviral target.

A structural and mechanistic understanding for how CoV N protein performs its myriad functional roles is limited. N proteins contain two RNA-binding domains of known structure: an N-terminal RNA binding domain (NTD) and a C-terminal dimerization domain (CTD)[20] linked by a Ser/Arg (SR)-rich linker (Fig. S1). Biochemical data suggest that the CTD is involved in oligomerization of N dimers,[20-22] and a small-angle x-ray scattering study suggests that the NTD and CTD do not interact in the absence of RNA.[23] The structures of the CTD from avian infectious bronchitis virus (IBV) and SARS-CoV reveal a tightly intertwined domain-swapped dimer[22,24] with the CTD N-terminal region, rich in basic amino acids, implicated in nucleic acid binding.[25,26] The structures of the NTD from SARS-CoV[26,27] and IBV[24,28] N have also been reported. The SR-rich region has been implicated in RNA binding in MHV[15] and in regulation of the oligomerization of SARS-CoV N[29] and a recent report provides genetic evidence for N-N interactions mediated by the NTD.[30] The reported affinity of N for U_{20} is in the 1-10 μM range (K_d) with no evidence for or against RNA binding specificity.[23]

Operating from the premise that the NTD and CTD fold independently into separable RNA binding domains,[23] we show here that the isolated NTD makes a specific, high affinity complex with the TRS and efficiently melts a TRS-cTRS duplex. These are two necessary features of a role in stimulating template switching during discontinuous sgRNA transcription. A mutation that cripples duplex TRS unwinding is defective in stimulation of CoV replication in cell culture; these studies suggest that specific targeting of the N NTD may lead to new antiviral agents.

Results

The MHV N NTD specifically binds to the TRS RNA with high affinity

Since it is known that N plays an important role in sgRNA synthesis[31] and can be crosslinked to the 5′ leader RNA in infected cells,[15,32] we hypothesized that N makes a high affinity interaction with the TRS, a highly conserved hexanucleotide sequence (Fig. 1). To test this, we measured the binding affinity of a 5′ fluorescein (F)-labeled decanucleotide corresponding to the MHV TRS (F-5′-gAAU-**CUAAAC**) with N219, an N domain protein fragment containing the folded NTD and the immediately adjacent intact SR-rich region (residues 60-219; Fig. S1), by fluorescence anisotropy. These data reveal that the N219-TRS complex is characterized by a $K_{\mathrm{obs}}=9.0\times10^{7}\ \mathrm{M}^{-1}$ at 150 mM K^{+}, 25 °C (Fig. 1, Table 1). To address the nucleotide specificity of NTD, we carried out fluorescence anisotropy-based RNA competition experiments with unlabeled mutant TRS RNAs (Fig. 1c; Table 1). Essentially all mutations in the TRS result in a decrease in K_{obs}, with a random RNA of the same length binding 53-fold less tightly. Substitution of ^{65}UCU^{67} with ^{65}GAG^{67} (TRS-Y3r) results in a modest ≈2-fold decrease in K_{obs} whereas complete replacement of the ^{68}AAA^{70} sequence with ^{68}CUU^{70} (TRS-R3y) results in a 20-fold decrease. Combining these two blocks of mutations into the same RNA (TRS-YR) suggests that these two effects are not additive ($\Delta G_c=-1.1$ kcal mol^{-1}), thus revealing that the ^{68}AAA^{70} to ^{68}CUU^{70} substitution is globally destabilizing to the interface. Finally, the complementary TRS sequence, 5′-AGUUUAGAUU (cTRS), adheres exactly to the 5′-RRYYYRRRYY motif present in the TRS; consistent with this, the affinity of N219 for a cTRS labeled with the rhodamine derivative, DY547, gives $K_{\mathrm{obs}}=9.1\times10^{7}\ \mathrm{M}^{-1}$ (Table 1). These data taken collectively reveal that the MHV NTD forms a specific, high affinity complex with both the TRS and cTRS RNA sequences that would be present in the leader and body TRSs and the nascent minus-strand RNA transcript, respectively, during sgRNA transcription.

Analysis of RNAs harboring successive 1-to 5-nucleotide deletions from the 5′ end of the TRS RNA decanucleotide suggest that these nucleotides upstream of the ^{68}AAA^{70} motif provide electrostatic stabilization to the complex, with $\Delta\Delta G_{\mathrm{obs}}$ per loss of successive phosphate groups as anticipated from a simple polyelectrolyte binding model[33] (Fig. 2 and Table S2).

The SR-rich region does not engage in specific interactions with the TRS RNA

Previous studies suggested that the SR-rich region provides most of the binding determinants for the specific interaction with the leader RNA in MHV.[15] To investigate the contribution of the SR-rich region

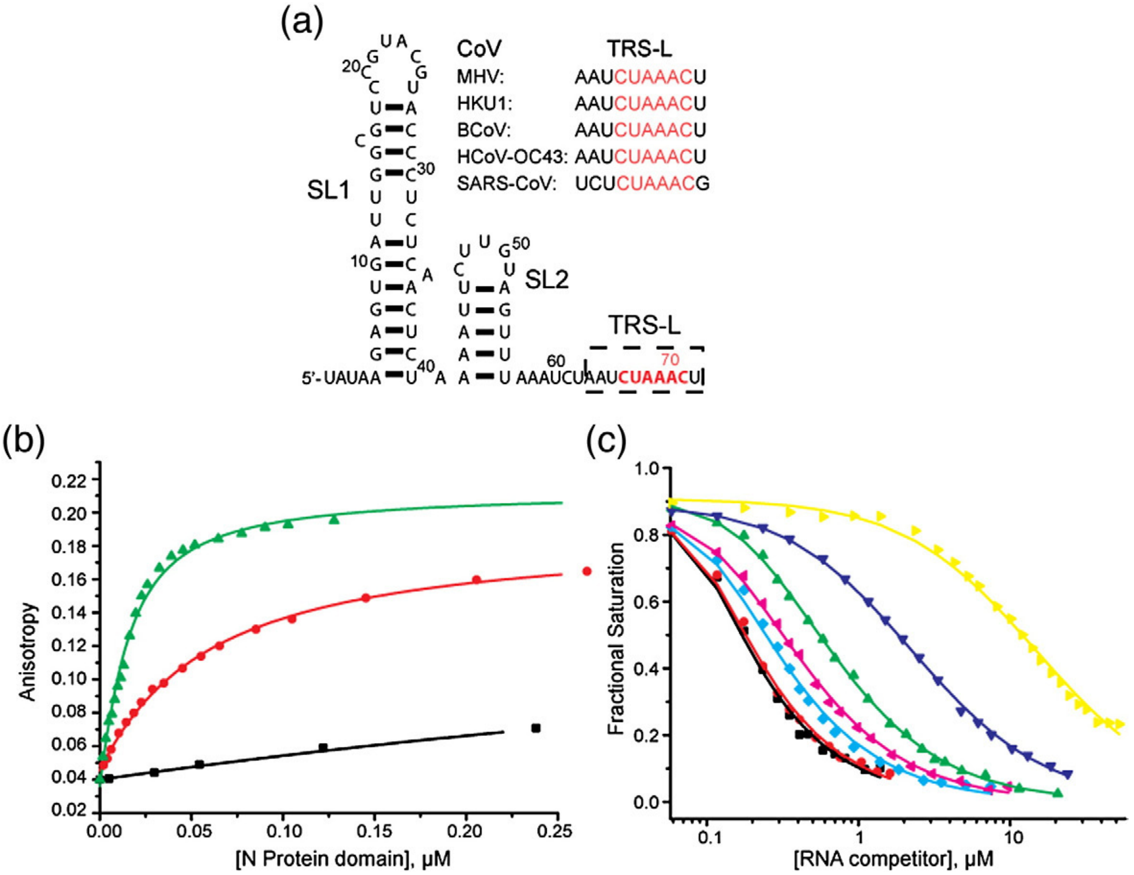

Fig. 1. MHV N NTD RNA binding assays. (a) The MHV 5′ leader sequence consisting of the first 72 nucleotides.[40] The 3′ most ten nucleotides containing the conserved hexanucleotide core TRS sequence (*red*) was used for binding assays. (b) Fluorescence anisotropy titrations of MHV N219 (*green*), N197 (*magenta*) and the CTD dimer (*black*) to 5′-fluorescein (F)-labeled TRS in 50 mM KP$_i$, pH 6.0, 100 mM KCl, 25 °C. The solid lines represent best fits to a one site (N219 and N197) or two site (CTD) binding model (fitting models (1) and (3), respectively, Materials and Methods). (c) Dissociation of an N219-5′-F-TRS complex by unlabeled mutant TRS RNAs as monitored by anisotropy. RNA competitors are unlabeled TRS (*black*), TRS-A69u (*red*), TRS-Y3r (*cyan*); TRS-A68u (*magenta*), TRS-A70u (*green*); TRS-R3y, (*blue*); random 9mer (*yellow*). The continuous lines through the data are defined by the binding parameters obtained by analysis using fitting model (2) (Materials and Methods) and are compiled in Table 1.

in TRS binding, we determined the affinity of N197, an NTD construct lacking the SR-rich region (residues 60-197; Fig. S1) (Fig. 1b), for TRS RNA ($K_{\text{obs}}=1.9 \pm 0.1\times10^7$ M^{-1}; Table 1). N197 makes a high affinity complex with TRS, but one characterized by an approximately 5-fold decrease when

Table 1. Binding affinities of MHV N protein domains for TRS RNAs[‡]

N Protein	RNA	Sequence (5′ to 3′)	*n*	Δ*H* (kcal mol^{-1})	K_{obs} (x10^7 M^{-1})	Fold decrease
N219	TRS	gAAUCUAAAC	1.1 ± 0.2	−23.4 ± 0.3	9.0 ± 0.9	–
N219	TRS-A70u	gAAUCUAAuC	1.0 ± 0.1	−15.3 ± 0.3	2.6 ± 0.4	3.5
N219	TRS-R3y	gAAUCUcuuC	nd*	nd	0.44 ± 0.01	20
N219	TRS-Y3r	gAAgagAAAC	nd	nd	4.1 ± 0.2	2.2
N219	TRS-YR	gAAgagcuuC	nd	nd	1.4 ± 0.7	6.4**
N219	TRS$^{\text{syn}}$	AAUCUAAACU	nd	nd	9.2 ± 0.8	1.0
N219	cTRS	AGUUUAGAUU	1.1 ± 0.2	−23.7 ± 0.1	9.1 ± 0.8	1.0
N219	cTRS-R3y$^{\text{syn}}$	AGUUUcuuUU	0.9 ± 0.1	−19.9 ± 0.3	0.54 ± 0.03	17
N219	Random10	auauagcuac	nd	nd	0.17 ± 0.09	53
N219	SARS-TRS	CUUCUCUAAACGAAG	1.1 ± 0.2	−15.2 ± 0.3	2.9 ± 0.8	3.1
N197	TRS	gAAUCUAAAC	nd	nd	1.86 ± 0.09	4.8
Y127A N219	TRS	gAAUCUAAAC	1.0 ± 0.2	−21.0 ± 0.1	0.48 ± 0.06	19
CTD dimer	TRS	gAAUCUAAAC	no***	no***	0.03 ± 0.01	300

[‡]Conditions: pH 6.0, 0.15 M K$^+$, 25.0 °C. K_{obs} values are averaged values from direct or competition fluorescence anisotropy-based titrations and isothermal titration calorimetry (ITC). *n* and Δ*H* are reported from ITC measurements. Other fitted parameters are summarized in Tables S1-S4. *nd, not determined. **The pairwise cooperative (coupling) free energy, ΔG_c was calculated from $\Delta G_c = \Delta\Delta G_{\text{TRS-YR}} - (\Delta\Delta G_{\text{TRS-R3y+}}\ \Delta\Delta G_{\text{TRS-Y3r}})$, where $\Delta\Delta G = \Delta G_{\text{mutant}} - \Delta G_{\text{TRS}}$. This gives $\Delta G_c = -1.1$ kcal mol^{-1}. ***no, no heat observed.

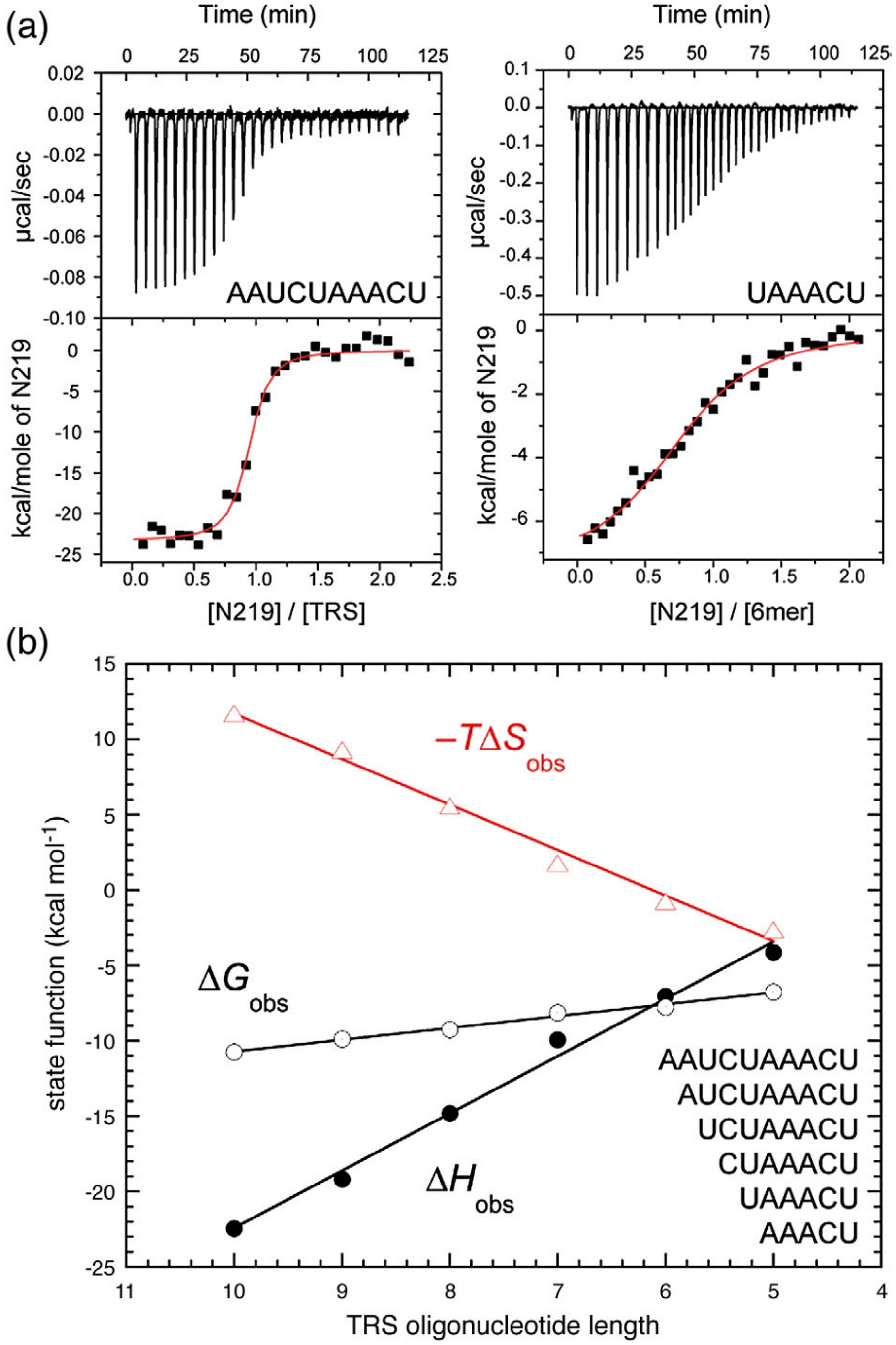

Fig. 2. (a*)* Representative ITC titrations of N219 into wild-type 10-mer TRS (1 µM) and a 5′ truncated 6-mer (5′-UAAACU; 25 µM) in 50 mM K^+phosphate, 100 mM KCl, pH 6.0 at 25 °C. The *red* line indicates the best fit according to a one-site binding model (see Table S2 for fitted parameters). (b) A *linear* dependence of all three state functions, ΔG_{obs}, ΔH_{obs}, and ΔS_{obs} is observed on the length of the 5′-hydroxylated TRS, with $\Delta\Delta G_{obs}$, $\Delta\Delta H_{obs}$, and $\Delta(-T\Delta S_{obs})$ of –0.8, –3.8 and +3.0 kcal mol^{-1} per nucleotide (or phosphodiester) deletion, respectively. The fact that nucleotide substitutions of the same 5′ bases are far less destabilizing (see Table 1), coupled with the same energetics penalty for loss of successive nucleotides (or –1 charge) from the chain strongly suggests that this region of the TRS stabilizes the N219 complex electrostatically, with the core ^{68}AAA^{70} motif providing significant nonelectrostatic stabilization to the complex.

compared to N219; this suggests that the nucleobase-specific interactions are contained entirely within the NTD. The increase in binding affinity is likely due to the presence of five additional positive charges from the SR-rich region, contributing a larger electrostatic component to the binding energy in N219 vs. N197. To test this, standard "salt-back" dissociation experiments were carried out to obtain information on the extent to which electrostatic interactions stabilize the NTD-TRS complex.[34,35] SK_{obs}, the dependence of K_{obs} on $[K^+]$, for the N219-TRS interaction is large (SK_{obs} = –5.5), consis-

Table 2. X-ray data collection and refinement statistics for MHV N197 (60-197)

	MHV N residues 64-194
Data Collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	34.16, 46.86, 71.71
Resolution	22.3–1.75 (1.78–1.75)*
I/σI	37.1 (2.9)
Completeness (%)	99.5 (97.2)
Redundancy	5.2 (3.7)
Refinement	
Resolution (Å)	22.3–1.75 (1.78–1.75)
R_{sym}	0.049 (0.405)
No. reflections	
Used for refinement	11,055
Used for R_{free} calculation	1,078
R factor (%)	18.5
R_{free} (%)	22.6
No. atoms	
Protein	1046
Water molecules	127
B-factors (Å^2, average)	
Protein	28.94
Water molecules	35.97
r.m.s. deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.064

* Values in parentheses are for the highest resolution shell.

tent with 7-8 ionic interactions in the complex for the RNA binding (Fig. S2), with 55% of the total binding free energy at 0.15 M K$^+$contributed by the polyelectrolyte effect. In contrast, the SK_{obs} of N197 is smaller, –3.9, with the polyelectrolyte contribution only ≈40% under these conditions. Thus, in this simplified polyelectrolyte model, N219 engages in 2-3 additional electrostatic interactions with the RNA, likely contributed by a subset of the C-terminal Arg residues in N219 vs. N197.

MHV NTD adopts a U-shaped β-platform structure

To begin to understand the molecular determinants of the interaction between the TRS and NTD, we solved the crystallographic structure of MHV N197 (residues 60-197), using the structure of the SARS NTD[36] as a search model for molecular replacement (Table 2). The structural model (Fig. 3) encompasses residues 64-194, with only the side chain of K113 in the β2′-β3′ hairpin loop modeled as an Ala due to poor side chain density. The 130-residue MHV NTD adopts a U-shaped β-platform which contains five short β-strands (arranged β4-β2-β3-β1-β5) across the platform and, as expected, adopts a fold that is nearly identical to NTDs of other coronaviral N proteins.[24,26-28,36] The putative RNA binding groove is characterized by the palm of the β-platform and an extended β-hairpin that collectively contain a large number of basic and aromatic amino acids that are proposed to directly interact with RNA (Fig. 3(a)-(b)). The base of the hairpin loop is strongly positively charged (Fig. 3(c)), with the temperature factors increasing as one moves away from the platform region to the tip of the β2′-β3′ hairpin (Fig. 3d). On the other hand, the C-terminal SR-rich region may effectively extend the RNA binding groove of N197 in N219.

Mutations in N219 influence the TRS binding affinity

We next determined the binding affinities of R110A, Y127A and Y129A N219s for the TRS RNA using our anisotropy-based assay. Although R110A and Y129A N219s each show only a modest decrease in binding affinity, characterization of the Y129A/R110A double mutant suggests that these the two residues are modestly energetically coupled (ΔG_c=–0.4 kcal mol^{-1}), consistent with a long-distance cooperativity across the β-platform (Table S3, Fig. S3). The difference in binding of this double mutant to the TRS-R3y RNA relative to N219 is identical to that observed for the wild-type TRS RNA; these data suggest that Y129 and R110 are unlikely to make base-specific contacts with-AAA-sequence (Table S3). In contrast, Y127A N219 binds the TRS with ≈ 19-fold decrease in affinity relative to wild-type N219. ^{1}H-^{15}N HSQC spectra of Y127A N219 suggest only localized structural perturbations in the mutant (Fig. S4).

TRS binding to N219 is strongly enthalpically driven

Given the anticipated involvement of aromatic residue-nucleobase stacking as an important part of the NTD-TRS interface, we next sought to understand the underlying thermodynamic origins of the binding affinity by ITC (Fig. 4 and Table S4). For both wild-type and Y127A N219s, complex formation is characterized by a significant enthalpic driving force (Table 1). Of particular note is that the difference in binding free energy between these two N219 proteins is entirely enthalpic in nature, *i.e.* $\Delta\Delta G=\Delta\Delta H$, with $\Delta(-T\Delta S)=0$. This is consistent with a direct π-stacking interaction between Y127 and one or more TRS nucleotides (Fig. 4) although other structural scenarios are possible. In contrast, the energetics of the binding of N219 to the TRS-A70u RNA, which harbors a single base substitution of the –^{68}AAA^{70}-motif, reveal a significant decrease in the entropic penalty coupled with a vastly different ΔH relative to the wild-type TRS RNA. This suggests a different mode of binding for this mutant RNA to N219 (Table 1).

MHV N NTD binds tightly to the SARS-CoV TRS

The high conservation of residues in the palm of the CoV NTDs[24,26-28] and the core TRS (Fig. 1) makes the prediction that the MHV NTD should form a non-cognate complex with the SARS-CoV TRS. Two putative TRS sequences have been proposed for SARS-CoV,[37,38] with the first containing the 5′-CUAA*AC* core observed in other CoVs and the second, 5′-*AC*GAAC, just downstream and

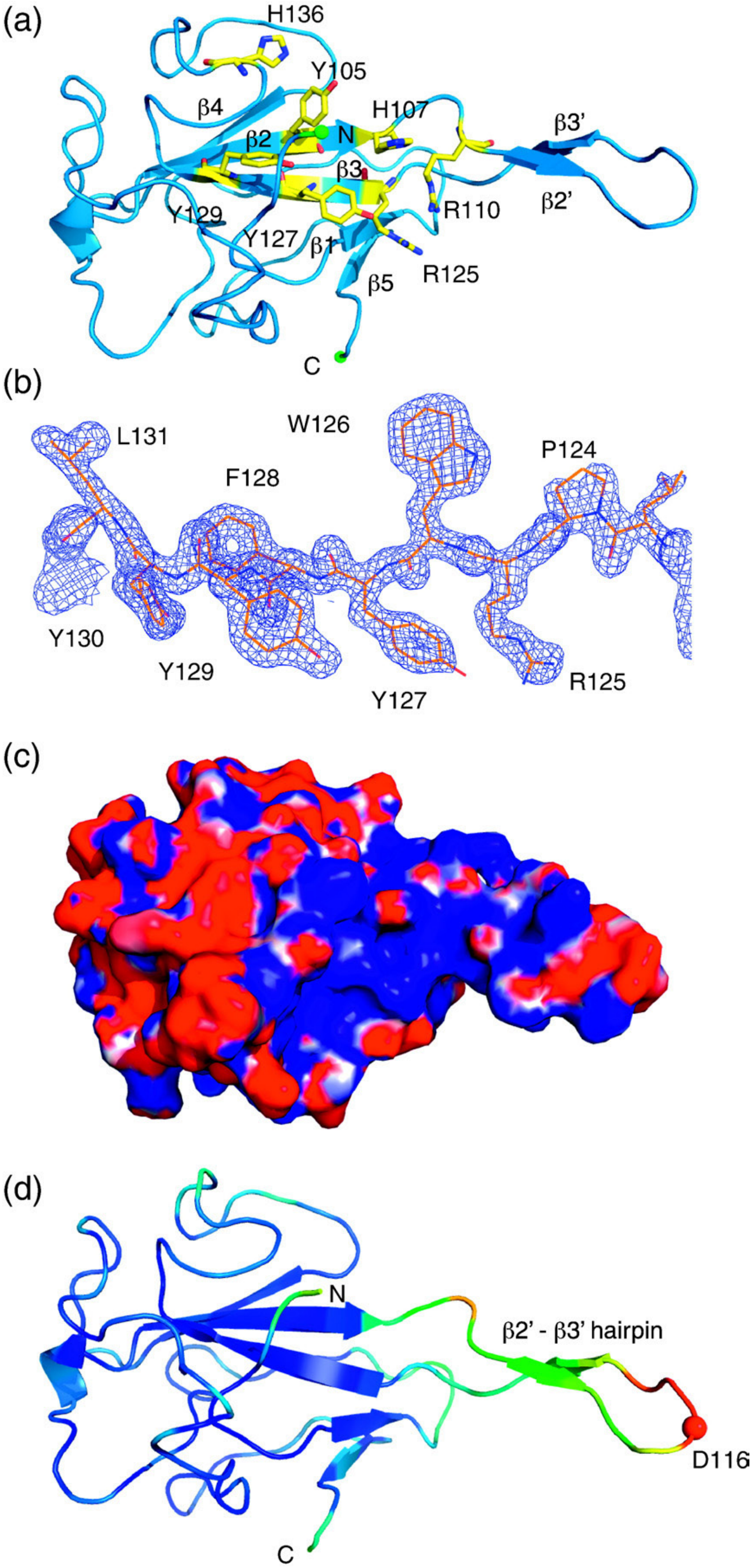

Fig. 3. Crystallographic structure of MHV N197. (a) Ribbon diagram of MHV NTD shown with candidate RNA binding residues (*yellow*). (b) The final refined $2mF_o$-dF_c electron density map of residues 124-131 is contoured at 1.5 σ to demonstrate the data quality of key β-platform residues in the model. (c) The electrostatic surface potential of MHV NTD shows the putative RNA binding region marked by a large electrostatic positive surface. (d) The ribbon diagram of MHV NTD shaded according to temperature factors of the Cα atoms. The color is ramped from dark blue (12.9-23.4 $Å^2$) to cyan (24.3-40.4 $Å^2$) to green (42.9-53.6 $Å^2$) to yellow (53.8-69 $Å^2$) to red (82.4-95.1 $Å^2$). The N-and C-termini are indicated, as is the Cα atom of D116 in the extended β2′-β3′ hairpin. Same view as in panels (a) and (c).

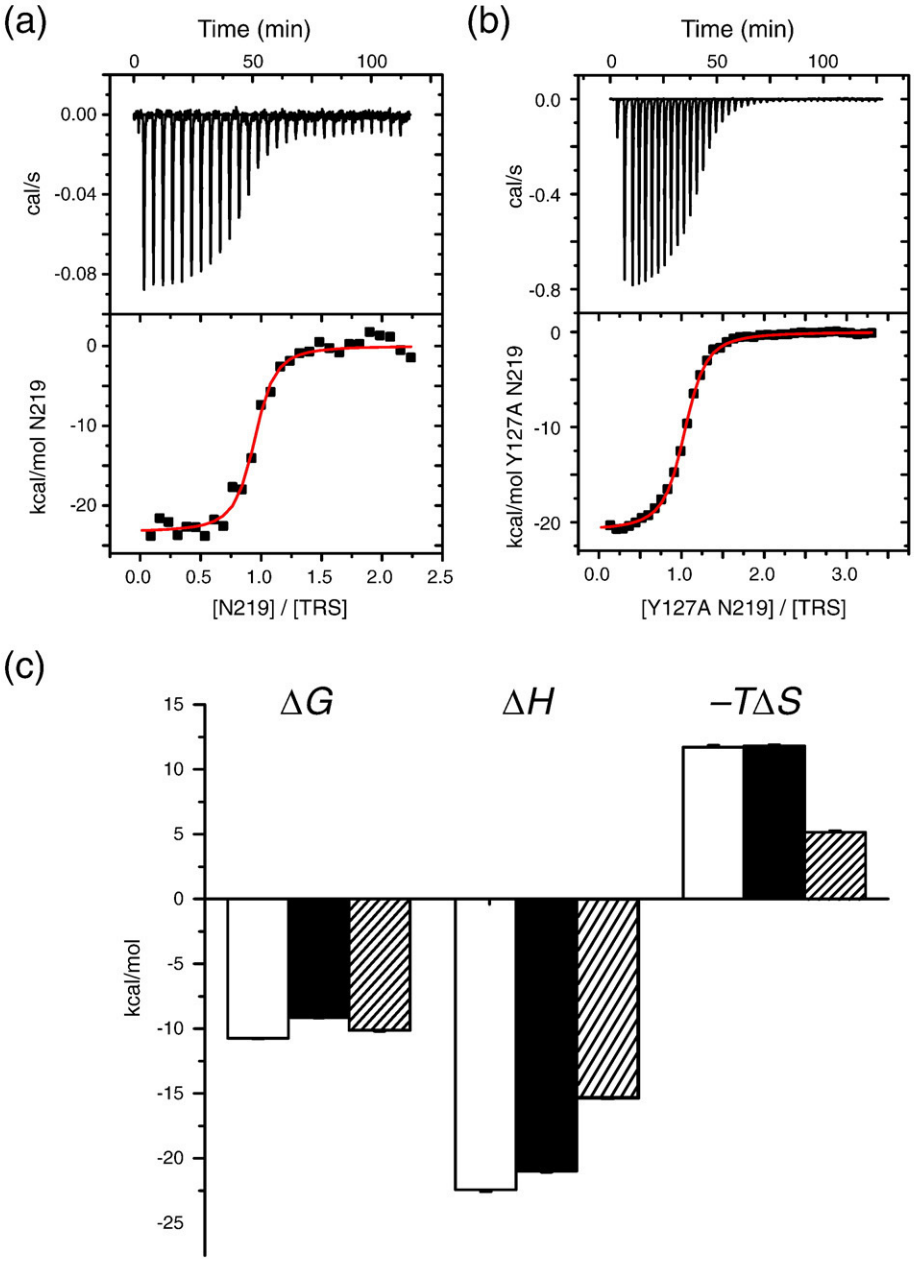

Fig. 4. Thermodynamic summary of MHV N NTD-TRS binding equilibria. Representative ITC titrations of (a) 10 μM N219 into 1 μM TRS and (b) 220 μM Y127A N219 into 10 μM TRS in 50 mM phosphate pH 6.0, 100 mM KCl and 25 °C. The *red* lines indicate the best fit to a one site binding model with the following parameters: N219, $n = 0.92$ (±0.01), $K_{obs} = 1.4$ $(\pm 0.3) \times 10^7$ M^{-1}, $\Delta H = -23.3$ (±0.4) kcal mol^{-1}; Y127A N219: $n = 1.02$ (±0.01), $K_{obs} = 5.8$ $(\pm 0.2) \times 10^6$ M^{-1}, $\Delta H = -20.9$ (±0.1) kcal mol^{-1}). (c) Graphical comparison of the enthalpic (ΔH) and entropic ($-T\Delta S$) components of the free energy (ΔG) for N219 binding to TRS (*open bars*), Y127A N219 binding to TRS (*black bars*) and N219 binding to TRS-A70u RNA (see Fig. 1a) (*hatched bars*).

overlapping the first (see Fig. 5a). Using the second putative TRS sequence, Baric and coworkers reported a 'rewiring' of the SARS-CoV genome by making parallel mutations in the TRS-L and TRS-B[39] (Fig. 5a); however, these mutations are not expected to appreciably affect the binding affinity of N for the SARS-CoV TRS. We tested this using a 15-nucleotide 5′-Cy3/3′-Cy5 labeled SARS-CoV TRS (Fig. 5b). By monitoring the anisotropy upon direct excitation of Cy5, we find that N219 binds to this RNA with a binding affinity of $K_{obs} = 2.9 \times 10^7$ M^{-1} (Table 1). The ≈3-fold decrease in affinity is explained by the fact that the SARS-CoV TRS may exist as a weak stem-loop,[40] giving rise to a competing equilibrium associated with melting the stem (Fig. 5a-b). The existence of the stem-loop in the doubly-labeled RNA was confirmed by a FRET efficiency (E) of ≈0.5 (Fig. 5c), a value consistent with a hairpin-unfolded RNA equilibrium (Fig. 5b). Regardless, stoichiometric N219 fully denatures this stem since E goes to zero. Companion ITC experiments further reveal that the ΔH_{obs}, is ≈8 kcal mol^{-1} less negative

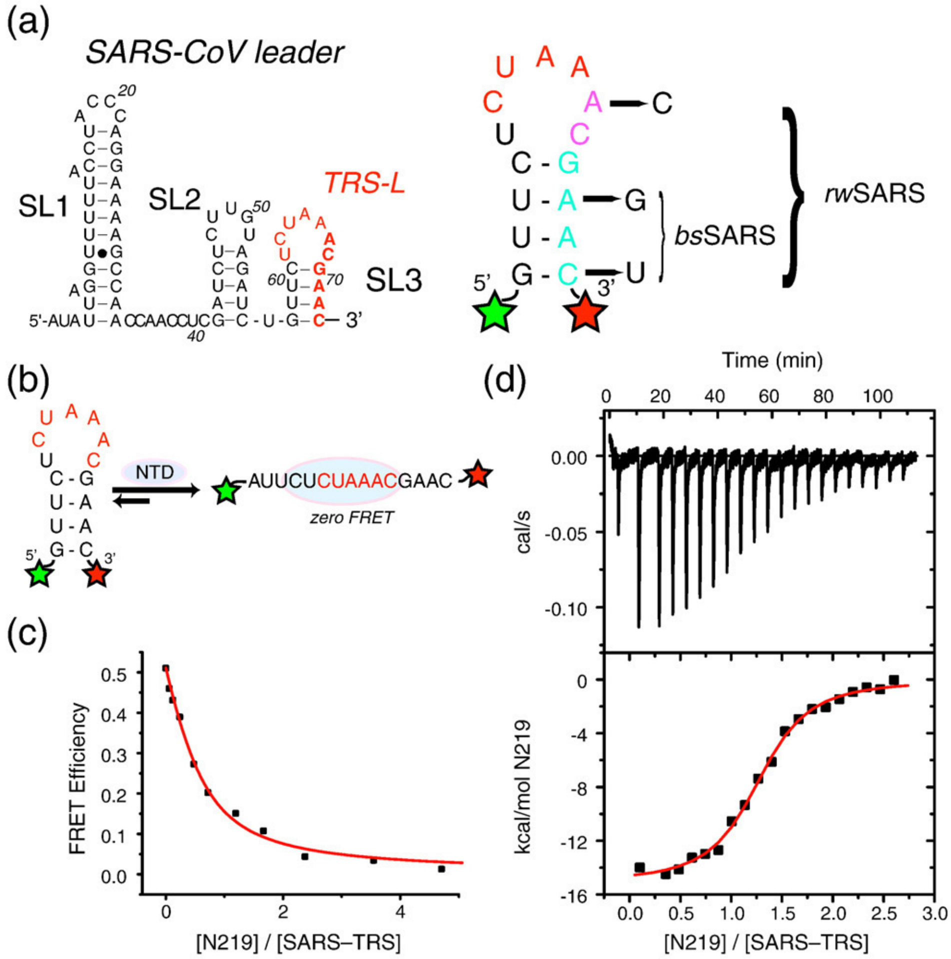

Fig. 5. MHV N219 binds to the noncognate SARS-CoV TRS RNA. (a) Schematic representation of SARS-CoV leader (*left*), with core (*red letters*) and alternative TRS (*blue letters; magenta letters*, overlapping region of the two TRSs highlighted. Broken stem and rewired SARS-CoV TRS sequences are also shown [39] (*right*). The *green* and *red* stars represent 5′ Cy3 and 3′ Cy5 FRET donor and acceptor dyes, respectively. (b) Cartoon representation of how N219 drives the hairpin-ssRNA equilibrium toward unfolded ssRNA. (c) Plot of FRET efficiency of the SARS-CoV TRS Cy3/Cy5 pair as a function of [N129]/[SARS-CoV TRS] ratio. The solid line represents a fit to a 1:1 binding model defined by the parameters in Table 1 as resolved by nonlinear least squares fit to model (4) (Materials and Methods). (d) Representative ITC titration of 37 μM N219 into 1 μM SARS TRS stem-loop labeled with a 5′-Cy3 and 3′-Cy5 FRET pair. Top panel, raw ITC data; bottom panel, concentration normalized enthalpy plotted vs. molar ratio. The red line depicts the best fit according to a one-site binding model where $n = 1.2\ (\pm 0.1)$, $K_{\mathrm{obs}} = 1.9\ (\pm 0.2) \times 10^7\ \mathrm{M}^{-1}$ and $\Delta H = -15.2\ (\pm 0.2)$ kcal/mol.

compared to the MHV TRS (Table 1, Fig. 5d, Table S4); this is as expected if endothermic stem melting is coupled to N219 binding. To verify this, a broken-stem mutant (bsSARS), which corresponds to two of the three mutations used in the rewiring study, was investigated along with the fully rewired TRS (rwSARS) RNA (Fig. 5a). The resulting increase in K_{obs} and $-\Delta H_{\mathrm{obs}}$ observed for each of these RNAs is consistent with N219-inducing melting of the helical stem in the wild-type SARS-TRS RNA (Table S4).

N219, but not Y127A N219 or N197, efficiently melts a duplex TRS

Since the N NTD makes a high affinity complex with both the TRS and cTRS, we hypothesized that it might melt an RNA duplex between the template TRS and nascent cTRS strand. We tested this using a FRET-based assay with a preformed 5′-Cy3-TRS–3′-Cy5-cTRS duplex RNA which is characterized by a FRET efficiency of ≈0.90 under these conditions (Fig. 6a,b). Addition of N219 results in an increase in the Cy3 emission intensity with a concomitant decrease in the Cy5 emission intensity to a FRET efficiency of zero, indicative of complete duplex melting. The subsequent addition of KCl to these mixtures results in dissociation of the N219-ssRNA complexes (see Fig. S2), and full recovery of the FRET efficiency associated with the duplex; this shows that N219-mediated RNA unwinding is fully reversible (Fig. S5). A quantitative analysis of these data to an equilibrium model that explicitly invokes the possibility that N219 binds to the duplex (K_4 in Model 5, Materials and Methods) reveals an affinity of<1 M^{-1}, *i.e.*, this complex does *not* form. In contrast, while the CTD dimer clearly binds to this duplex, it

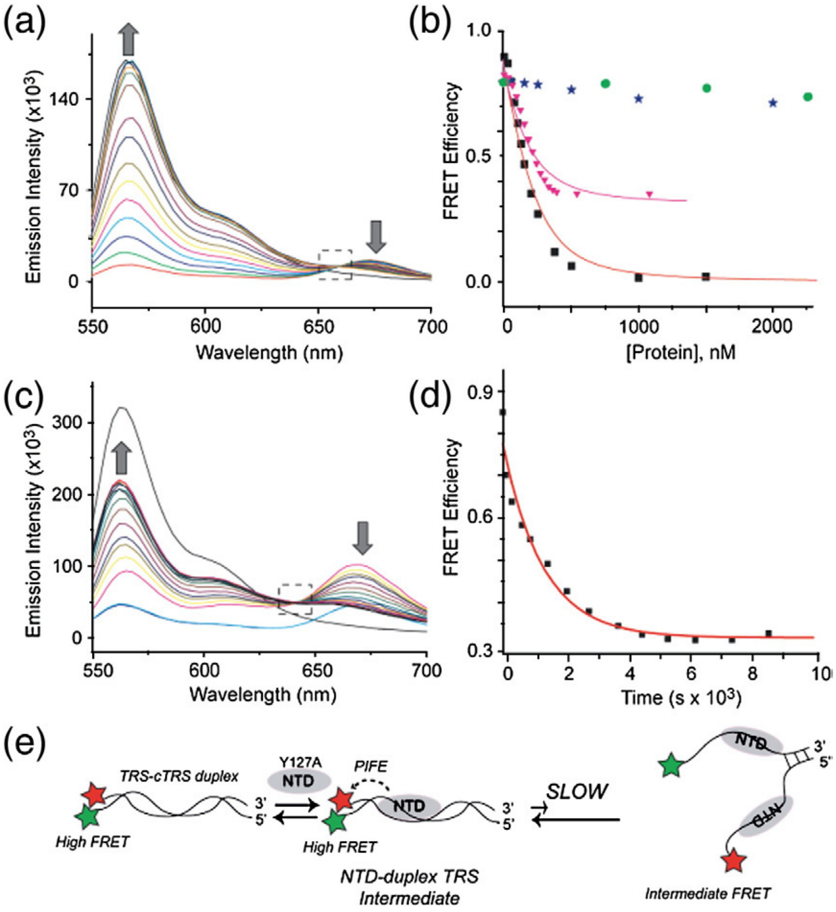

Fig. 6. N219 melts a TRS-cTRS duplex. (a) Representative fluorescence emission spectra of 5′-Cy3-TRS/3′-Cy5-rTRS RNA duplex obtained on titration with N219. Arrows, changes in the spectra as N219 is added; dashed box, isosbestic point. (b) Plot of FRET *E* as a function of total protein concentration. N219, *black symbols*; Y127A N219, *green*; N197, *magenta* and CTD dimer, *blue*. The *red* and *magenta* continuous lines define a best fit to a model that assumes that the NTD in each case forms a 1:1 complex with the component single strands with the affinities fixed to their known values (Table 1), and the affinity for the duplex RNA optimized in the fit (see Materials and Methods, model (5)). (c) Emission spectrum of 50 nM 5′-Cy3-TRS/3′-Cy5-rTRS dsRNA duplex (*cyan line*) upon addition of 12 μM Y127A N219 (*magenta line*) acquired at different time points. (d) Plot of FRET efficiency as a function of time from the data in panel (c). The continuous line is a best fit to a single exponential dissociation model with $k = 8\ (\pm1)\times10^{-4}\ \mathrm{s}^{-1}$. (e) Cartoon model of NTD-mediated unfolding of the TRS-cTRS duplex by Y127A N219.

is unable to denature it, even under conditions where K_{obs} are comparable for the two domains (35 mM K^+vs. 150 mM K^+) (Fig. 6b; Table S1).

Strikingly, while Y127A N219 is capable of melting the TRS-cTRS duplex, it is strongly kinetically impaired (Fig. 6c,d). In addition, there is a significant enhancement of the Cy5 emission intensity upon addition of Y127A N219 not observed with wild-type N219; this appears to be the result of a Y127A:dsRNA complex, since direct excitation of the single-stranded 3′-Cy5-cTRS RNA:Y127A N219 complex yields no such enhancement (Fig. S6). We interpret this as a ribonucleoprotein complex-mediated modulation of the environment of Cy5.[41] At a saturating concentration of Y127A N219, we observe a rate constant of $k = 8 \pm 1\times10^{-4}\ \mathrm{s}^{-1}$ or ≥30-fold slower than wild-type N219. Finally, for both Y127A N219 and WT N197, the FRET efficiency fails to return to zero even after very long incubation times with saturating N protein, as expected for the full duplex dissociation observed for wild-type N219 (Fig. 6b,d). This finding suggests an incomplete melting of the dsTRS RNA, implying a partially melted, long-lived intermediate complex with these two proteins (Fig. 6e). These observations reveal that the SR-rich tail and key residues on the β-platform, *e.g.*, Y127, function cooperatively to melt the duplex TRS in a kinetically facile manner.

Recovery and functional analysis of Y127A and Y129A N-containing viruses

To test the functional importance of these Tyr substitutions on viral replication, we electroporated BHK-R cells with Y127A and Y129A N-containing MHV genomes in the absence or presence of a "helper" RNA encoding a wild-type or mutant N gene. From electroporations of the Y127A N-containing MHV RNA, we were only able to recover (18/18) wild-type N genes from plaque-purified virion particles, irrespective of whether we used wild-type, Y127A, or no helper RNA. In contrast, Y129A N viruses were recovered regardless of the presence of the helper RNA.

We next tested the effect of these mutations in an infectious center assay. Here, cDNAs representing the wild-type MHV genome were *in vitro* transcribed and electroporated into BHK-R cells in the presence of either a WT, Y127A or Y129A N helper RNA, with the electroporated cells plated on confluent L2 cells, incubated at 37 °C for 5 h to allow for cell attachment, and then overlaid with agarose-containing media. Plaques were counted 3 days later (Fig. 7). Although we find a modest decrease in infectious center formation with the Y129A N helper RNA relative to the wild type helper, the number of infectious centers formed with the Y127A N helper RNA is identical to that observed in the absence of

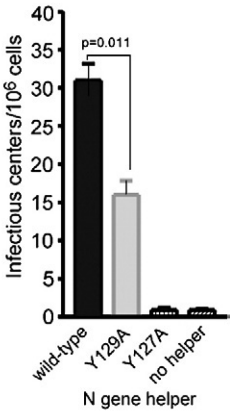

Fig. 7. The replication accessory function of N is severely impaired in the Y127A N NTD mutant relative to the Y129A and wild-type N. BHK-R cells electroporated with wild-type MHV genomic RNA in the presence of the indicated helper RNA were grown to confluence and infectious centers in L2 cells counted (mean and s.d. from three experiments).

any N helper RNA. These experiments taken collectively provide genetic evidence that the Y127A N-protein containing virions are strongly selected against, and that the NTD, and specifically Y127, plays a critical role in stimulation of viral replication.

Discussion

Although the atomic resolution structures of the N protein NTD and CTD from several coronaviruses are now available[22,24-28,36] (this work), detailed knowledge of the RNA binding properties of N is rather limited. We show here that MHV N219 forms a high-affinity complex with both the MHV and non-cognate SARS-CoV TRS[39] a finding that speaks to the conservation of the NTD-TRS interaction as a conserved feature of CoV replication, despite the distinct structural contexts of the leader TRS in each case. We also show that the N NTD possesses potent helix-destabilizing activity. The NTD employs enthalpically stabilizing base stacking interactions to drive high affinity and sequence-specific complex formation with the single-stranded TRS and cTRS RNAs. This binding, in turn, strongly enhances the rate of TRS-cTRS duplex melting that models an intermediate in sgRNA transcription by the coronaviral replicase complex. Full helix-destabilizing activity of the N NTD requires determinants on both the β-platform, *i.e.*, Y127, as well as the SR-rich domain. Although we have not determined the RNA binding specificity ratio of N197 (K_{sp}; see below), comparative studies of N197 and N219 suggest that the β-platform domain provides key specificity determinants for TRS recognition, with the SR-rich region stabilizing the complex via non-specific electrostatic interactions, likely with the region just 5′ to the TRS core sequence. These findings suggest that phosphorylation of S170 or T177 in MHV N,[10] the latter of which is in close proximity to the C-terminal β5 strand, or Ser residues within SR-rich domain,[9] might strongly modulate the ssRNA and helix-melting properties of the N protein.

The degree to which N219 is capable of discriminating between a short TRS-containing oligonucleotide vs. two "random" sequence RNAs of different base compositions, defined by the specificity ratio, $K_{sp} = K_{TRS}/K_{random}$, is ≈53 and ≈25 for a 10-mer and 9-mer RNAs, respectively. While K_{sp} is modest when compared to *bona fide* sequence-specific RNA binding proteins,[42] it appears to be of the same order of magnitude determined for another viral nucleocapsid protein, from HIV-1.[43] Such a relatively small specificity ratio is not inconsistent with a N219-TRS binding mode that is characterized by a sizable electrostatic contribution to the binding energy, as well as the multiple functional roles N protein must play in the viral life cycle. It is striking, nonetheless, that a single Y127A substitution within the highly conserved WY^{127}FY^{129}Y sequence on the β3-strand, like that of a complete pyrimidine substitution of the triple adenosine motif ^{68}AAA^{70} (in the 10-mer context), reduces the binding affinity to just three-fold above what we operationally define as non-sequence specific binding. In contrast, single adenosine to pyrimidine substitutions within this ^{68}AAA^{70} are not as destabilizing, but nearly additive ($\Delta\Delta G = 1.5$ kcal mol^{-1}) relative to the complete ^{68}AAA^{70} to ^{68}CUU^{70} substitution ($\Delta\Delta G = 1.8$ kcal mol^{-1}), with the 3′-most A70 making the largest single contribution to the N219-TRS binding energy. These data are consistent with a model in which the 3′ end of the TRS is anchored on the β-platform via enthalpically stabilizing aromatic base-stacking interactions with the 5′ side of the TRS held in place largely by electrostatic interactions that extend into the SR-rich tail.

R110, Y127, and Y129 form a nearly contiguous surface on the β-platform, with the central residue, Y127, functioning as a linchpin in what appears to be a cooperative unit with the sum of the $\Delta\Delta G$s for any two single mutations (*e.g.*, Y127A and Y129A; Y127A and R110A) greater than that observed for the corresponding double mutant. Interestingly, an Ala substitution of the residue analogous to Y127 in IBV N, Y92A, leads to a profound reduction in viral replication.[44] The ability of N219 to facilitate the melting of a duplex TRS may underscore the ability of N to stimulate template switching during sgRNA transcription, as well as function as a nucleic acid chaperone.[18] Our functional characterization of Y127A N reveals that that this substitution abrogates the ability of N to stimulate RNA replication, and the molecular origin of this defect is likely attributable to the kinetically crippled helix-unwinding activity of Y127A N.

It is also known that coronavirus N encapsidates viral RNA into ribonucleoprotein (RNP) particles[45] and SARS-CoV N has been implicated in playing an essential role in viral RNA packaging;[46] however, the mechanism of RNA packaging is far from clear.

A recent structural study of the SARS-CoV CTD led the authors to speculate that the CTD plays a key role in the helical nucleocapsid[11] assembly.[25,26] Our findings further suggest that the ability of the NTD to melt dsRNA may also play a role in RNA packaging or other steps of the viral life cycle where RNA remodeling is required.

A model for how N protein-catalyzed unwinding of a transiently formed dsRNA between the body TRS and the cTRS in the nascent (daughter) strand might stimulate template switching during subgenomic RNA transcription is shown in Fig. 8. In this model, template switching is an ordered unfolding of the TRS-cTRS duplex and subsequent hybridization of the nascent strand with the 5′ leader RNA. Biological studies in TGEV and SARS-CoV reveal that one or two nucleotides 5′ to the TRS core sequence, the core TRS itself, and ≤5 downstream nucleotides on the template strand are required to be identical to those in the leader TRS region for efficient sgRNA synthesis to occur.[7,39] This would optimally position key NTD recognition determinants, *e.g.*, the triple adenosine motif, in the TRS and cTRS RNAs in the middle of a TRS-cTRS duplex that likely forms behind the elongating RdRp.[47-49] NTD-mediated local unfolding here would lead to an increase in the lifetime of the nascent strand in an unpaired state, thus accelerating the rate of nucleation of base pairing with core leader TRS, allowing the RdRp to switch RNA templates.

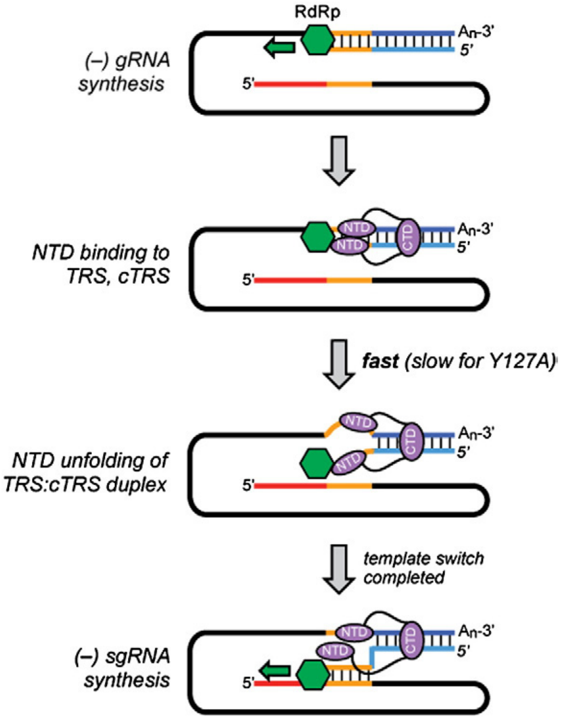

Fig. 8. Proposed model for the role of N protein in subgenomic RNA transcription. The transcriptase-replicase complex (labeled RdRp, *green* hexagon) initiates transcription from the 3′-polyadenylated end of the genomic RNA, and can either synthesize full-length (–) genomic RNA (*upper* panel), or switch templates to the 5′ leader at the TRS to synthesize (–) sgRNA by copying the remainder of the 5′ leader (*red* line) (*bottom* panel).[7] In the presence of an N protein dimer (*purple* ellipsoids connected by the *black* SR-rich linker), the two NTD domains bind to the TRS and cTRS sequences (*orange* lines), melt the duplex, and accelerate the rate of base pairing of the nascent cTRS with the leader TRS (*middle two* panels). The duplex RNA-binding activity of the CTD dimer may help anchor N to the nascent duplex.

It is not yet known if N interacts directly with any component of the polymerase complex, although antibodies against N strongly inhibit RNA transcription and N strongly stimulates virus replication.[16,50] If accumulating structural and biochemical evidence for a closed-to-open conformational switch enabling processive elongation by viral RdRps[47,49] characterizes coronaviral RdRps as well, a direct interaction with N might inhibit elongation and perhaps pause the polymerase complex just past the 5′ end of the TRS, providing time for a template switch to occur. Efficient reconstitution of an active coronaviral RdRp complex on defined RNA templates[51] will be required to test this model.

Materials and Methods

Preparation of RNA samples

Unlabeled TRS and TRS mutant RNAs were obtained by *in vitro* runoff transcription using SP6 RNA polymerase and purified by denaturing PAGE essentially as previously described.[52] This protocol necessitates the addition of non-native 5′-terminal guanosine residue to some RNAs, denoted by the lower-case "g". All other unlabeled or fluorescently labeled RNAs were obtained from Dharmacon or IDT and purified by denaturing PAGE.

Plasmid construction and protein expression and purification

For the plasmids encoding various fragments of MHV-A59 nucleocapsid protein, the coding sequences were amplified from the full-length MHV N gene using standard PCR based approaches. The PCR products were digested by NdeI and BamHI and ligated into pET3a, pET15b, or pGST-parallel expression plasmids.[53] The plasmids encoding the substitution mutants were prepared using QuickChange PCR-based mutagenesis of the wild type N219 overexpression plasmid as a template. The integrity of all the constructs was confirmed by DNA sequencing. Recombinant proteins were expressed from their respective pET3a-N197 (residues 60-197), pET15b-N219 (residues 60-219) and pGST-CTD (residues 256-385) plasmids, in *E. coli* BL21(DE3)/pLysS. The growth, expression and purification of N fragments expressed from pET3a, pET15a or pGST were carried out using the procedures described previously[54]; GST-CTD proteins had 2 mM DTT in the buffer throughout purification. The GST tag was cleaved from the CTD by TEV protease overnight at 4 °C in 50 mM Tris, 100 mM NaCl at pH 8.0. CTD was separated using cation exchange and further purified using a Superdex G75 chromatography; the retention time was consistent with that of a dimer (≈28 kDa). The protein purity by inspection of Coomassie-stained 18% Tris-glycine SDS-PAGE gels was estimated to be>95%. All proteins were further characterized by MALDI-TOF mass spectrometry. The concentration of purified proteins was determined using the calculated molar extinction coefficient at 280 nm, with proteins stored at –80 °C in concentrated aliquots.

Crystallization and structure determination

N197 overexpressed in *E. coli* and purified as described for N219 was concentrated to approximately 200 μM and buffered with 50 mM potassium phosphate (pH 6.0), 100 mM KCl. Crystals were grown via hanging-drop vapor diffusion against 30% PEG 1000, 50 μM CAPSO (pH 9.0) at 20 °C. Crystals grew overnight and were frozen in the well solution, 30% PEG 1000, 50 μM CAPSO (pH 9.0). Diffraction data were collected at-160 °C on an R-AXIS IV+ detector at Indiana University. The space group of the crystal was primitive orthrombic ($P2_12_12_1$) with one protein monomer in the asymmetric unit. Diffraction data to 1.75 Å was reduced using HKL-2000. Initial phases were determined using a portion of the crystal structure of SARS-CoV NTD (PDB: 2ofz) as a molecular replacement search model in phaser.[36] Iterative rounds of model building and refinement were carried out in Coot and Phenix, respectively. The N protein was then divided into ten segments by the TLSMD server[55] for TLS refinement. The quality of the final structure was verified using MOLPROBITY. A Ramachandran plot analysis revealed that 96.2% of residues are in the most favored regions and the remaining 3.8% of residues are found in additional allowed regions; no residues were found in disallowed regions. All structure-related figures were prepared using PyMOL (DeLano Scientific).

Fluorescence anisotropy and fluorescence resonance energy transfer (FRET) experiments

These experiments were typically performed on an ISS PC1 spectrofluorometer using 5.0 or 10.0 nM RNA (anisotropy) or 50 nM (FRET) RNA in 50 mM potassium phosphate, 100 mM KCl, pH 6.0, unless otherwise noted. TRS binding by N variants was measured by monitoring the change in the anisotropy of the labeled TRS. The binding of N variants to the unlabeled TRS and TRS mutant RNAs was followed using a standard competition assay. FRET experiments were carried out with RNA labeled with a Cy3-Cy5 pair ($\lambda_{ex}=520$ nm; $\lambda_{em}=550$-700 nm; Cy3 $\lambda_{max}=570$ nm; Cy5 $\lambda_{max}=670$ nm).with the FRET efficiency, E, calculated from $E=1-(I_{DA}/I_D)$, where I_D is the Cy3 quantum yield and I_{DA} is the Cy3 quantum yield in the presence Cy5, following a 2-10 min equilibration upon addition of the titrant. No change in the fluorescence intensity (quantum yield) of the component Cy3-and Cy5-labeled TRS and cTRS single-stranded RNAs, respectively, was observed; thus changes in I_i are directly attributed to FRET or protein-induced fluorescence enhancement (PIFE) (see Fig. S6).[41] Nonlinear least-squares fits to all binding isotherms were carried out using DynaFit[56] with the appropriate binding model (models (1)–(5), as indicated below).

Nonlinear Least Squares Fitting Models

Model (1): Equilibrium titration of N219 into fluorescently labeled RNAs

N219+ RNA∗ ⇆ N219-RNA* K_{obs}

Model (2): Competition equilibrium titration of unlabeled mutant RNAs into N219-labeled RNA complex

N219+ RNA* ⇆ N219-RNA* K_1 (fixed)
(by direct titration with Model (1))
N219-RNA*+ RNA ⇆ RNA*+ N219-RNA K_{obs}

Model (3): Sequential 2-site equilibrium titration of the CTD dimer (CTD) into labeled TRS

CTD+ RNA ⇆ CTD-RNA K_1
CTD+ CTD-RNA ⇆ (CTD)$_2$-RNA K_2

Model (4): Equilibrium titration of N219 into a double-labeled SARS-TRS RNA hairpin using FRET

dSARS-TRS ⇆ SARS-TRS (stem SL3 unfolding) K_1(fixed by $E=0.5$)
N219+ SARS-TRS ⇆ N219-SARS-TRS K_{obs}

Model (5): Equilibrium titration of NTDs into a duplex TRS-cTRS FRET pair

TRS+ cTRS ⇆ dsTRS (TRS annealing) K_1 (fixed)
N219+TRS ⇆ N219-TRS K_2 (fixed)
N219+ cTRS ⇆ N219-cTRS K_3 (fixed)
N219+ dsTRS ⇆ N219-dsTRS K_4

Isothermal Titration Calorimetry

Isothermal titration calorimetry experiments were carried out using a MicroCal VP-ITC calorimeter. In a typical experiment, 20 μM protein was titrated into 1 μM RNA in 50 mM K$^+$phosphate, pH 6.0 and 100 mM KCl at 25.0 °C, unless otherwise noted. All experiments were carried out in triplicate and the averaged values reported. Best fits were generated using a single site binding model described previously.[57]

Recovery and Characterization of Mutant Viruses

The cDNA *in vitro* assembly reverse genetic system described previously[58] was used to generate viral genomes containing the N Y127A and Y129A mutations. To generate mutant viruses, cDNAs representing the entire MHV genome were constructed by sequential ligation of the A-G cDNA fragments as described previously.[58,59] The ligated cDNAs representing mutant or wild type N gene-containing MHV genomes were *in vitro* transcribed and electroporated into BHK-R cells in the presence of a wild-type or mutant N gene transcript as previously described.[40] Cultures were observed for up to 72 h for the development of cytopathic effect (cell fusion) and harvested by freezing at –70 °C. The recovered viruses were plaque isolated and expanded on DBT cells. Total RNAs were extracted using QIAGEN RNeasy kit. The entire N gene of each plaque isolate as well as their 5′ and 3′ UTRs were sequenced to verify the genotype of the recovered viruses.

Accession codes

Atomic coordinates and structure factors of MHV N NTD have been deposited in the Protein Data Bank under accession code 3hd4.

Acknowledgements

The authors acknowledge support from R01 AI067416 (to D. P. G. and J. L. L.) and R01 AI040187 (to D. P. G.) from the NIH.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.09.040

References

1. Zhong, N. S., Zheng, B. J., Li, Y. M., Poon, L. L. M., Xie, Z. H., Chan, K. H. *et al.* (2003). Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *The Lancet*, **362**, 1353–1358.
2. Pyrc, K., Jebbink, M. F., Berkhout, B. & van der Hoek, L. (2004). Genome structure and transcriptional regulation of human coronavirus NL63. *Virol. J.* **1**, 7.
3. Woo, P. C., Lau, S. K., Chu, C. M., Chan, K. H., Tsoi, H. W., Huang, Y. *et al.* (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* **79**, 884–895.
4. De Albuquerque, N., Baig, E., Ma, X., Zhang, J., He, W., Rowe, A. *et al.* (2006). Murine hepatitis virus strain 1 produces a clinically relevant model of severe acute respiratory syndrome in A/J mice. *J. Virol.* **80**, 10382–10394.
5. Pasternak, A. O., Spaan, W. J. M. & Snijder, E. J. (2006). Nidovirus transcription: how to make sense...? *J. Gen. Virol.* **87**, 1403–1421.
6. Sawicki, S. G. & Sawicki, D. L. (1998). A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* **440**, 215–219.
7. Zuniga, S., Sola, I., Alonso, S. & Enjuanes, L. (2004). Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.* **78**, 980–994.
8. Li, L., Kang, H., Liu, P., Makkinje, N., Williamson, S. T., Leibowitz, J. L. *et al.* (2008). Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. *J. Mol. Biol.* **377**, 790–803.
9. Calvo, E., Escors, D., Lopez, J. A., Gonzalez, J. M., Alvarez, A., Arza, E. *et al.* (2005). Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells. *J. Gen. Virol.* **86**, 2255–2267.
10. White, T. C., Yi, Z. & Hogue, B. G. (2007). Identification of mouse hepatitis coronavirus A59 nucleocapsid protein phosphorylation sites. *Virus Res.* **126**, 139–148.
11. Barcena, M., Oostergetel, G. T., Bartelink, W., Faas, F. G., Verkleij, A., Rottier, P. J. *et al.* (2009). Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. *Proc. Natl Acad. Sci. USA*, **106**, 582–587.
12. Baric, R. S., Nelson, G. W., Fleming, J. O., Deans, R. J., Keck, J. G., Casteel, N. *et al.* (1988). Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. *J. Virol.* **62**, 4280–4287.
13. Eleouet, J. F., Slee, E. A., Saurini, F., Castagne, N., Poncet, D., Garrido, C. *et al.* (2000). The viral nucleocapsid protein of transmissible gastroenteritis coronavirus (TGEV) is cleaved by caspase-6 and-7 during TGEV-induced apoptosis. *J. Virol.* **74**, 3975–3983.
14. Stohlman, S. A., Baric, R. S., Nelson, G. N., Soe, L. H., Welter, L. M. & Deans, R. J. (1988). Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J. Virol.* **62**, 4288–4295.
15. Nelson, G. W., Stohlman, S. A. & Tahara, S. M. (2000). High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. *J. Gen. Virol.* **81**, 181–188.
16. Compton, S. R., Rogers, D. B., Holmes, K. V., Fertsch, D., Remenick, J. & McGowan, J. J. (1987). In vitro replication of mouse hepatitis virus strain A59. *J. Virol.* **61**, 1814–1820.
17. Thiel, V., Ivanov, K. A., Putics, A., Hertzig, T., Schelle, B., Bayer, S. *et al.* (2003). Mechanisms and enzymes involved in SARS coronavirus genome expression. *J. Gen. Virol.* **84**, 2305–2315.
18. Zúñiga, S., Sola, I., Moreno, J. L., Sabella, P., Plana-Durán, J. & Enjuanes, L. (2007). Coronavirus nucleocapsid protein is an RNA chaperone. *Virology*, **357**, 215–227.
19. Huang, M., Maynard, A., Turpin, J. A., Graham, L., Janini, G. M., Covell, D. G. *et al.* (1998). Anti-HIV agents that selectively target retroviral nucleocapsid protein zinc fingers without affecting cellular zinc finger proteins. *J. Med. Chem.* **41**, 1371–1381.
20. Chang, C.-k., Sue, S.-C., Yu, T.-h., Hsieh, C.-M., Tsai, C.-K., Chiang, Y.-C. *et al.* (2006). Modular organization of SARS coronavirus nucleocapsid protein. *J. Biomed. Sci.* **13**, 59–72.
21. Surjit, M., Liu, B., Kumar, P., Chow, V. T. K. & Lal, S. K. (2004). The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain. *Biochem. Biophys. Res. Commun.* **317**, 1030–1036.
22. Yu, I.-M., Oldham, M. L., Zhang, J. & Chen, J. (2006). Crystal Structure of the Severe Acute Respiratory Syndrome (SARS) Coronavirus nucleocapsid protein dimerization domain reveals evolutionary linkage between Corona-and Arteriviridae. *J. Biol. Chem.* **281**, 17134–17139.
23. Chang, C.-K., Hsu, Y.-L., Chang, Y.-H., Chao, F.-A., Wu, M.-C., Huang, Y.-S. *et al.* (2009). multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: Implications for ribonucleocapsid protein packaging. *J. Virol.* **83**, 2255–2264.
24. Jayaram, H., Fan, H., Bowman, B. R., Ooi, A., Jayaram, J., Collisson, E. W. *et al.* (2006). X-ray structures of the N-and C-terminal domains of a coronavirus nucleocapsid protein: implications for nucleocapsid formation. *J. Virol.* **80**, 6612–6620.
25. Takeda, M., Chang, C.-k., Ikeya, T., Güntert, P., Chang, Y.-h., Hsu, Y.-l. *et al.* (2008). Solution structure of the C-terminal dimerization domain of SARS Coronavirus nucleocapsid proteins solved by the SAIL-NMR method. *J. Mol. Biol.* **380**, 608–622.
26. Chen, C.-Y., Chang, C.-k., Chang, Y.-W., Sue, S.-C., Bai, H.-I., Riang, L. *et al.* (2007). Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. *J. Mol. Biol.* **368**, 1075–1086.
27. Huang, Q., Yu, L., Petros, A. M., Gunasekera, A., Liu, Z., Xu, N. *et al.* (2004). Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry*, **43**, 6059–6063.
28. Fan, H., Ooi, A., Tan, Y. W., Wang, S., Fang, S., Liu, D. X. *et al.* (2005). The nucleocapsid protein of coronavirus infectious bronchitis virus: Crystal structure of its N-terminal domain and multimerization properties. *Structure*, **13**, 1859–1868.
29. Luo, H., Ye, F., Chen, K., Shen, X. & Jiang, H. (2005). SR-Rich motif plays a pivotal role in recombinant SARS Coronavirus nucleocapsid protein multimerization. *Biochemistry*, **44**, 15351–15358.

30. Hurst, K. R., Koetzner, C. A. & Masters, P. S. (2009). Identification of in vivo-interacting domains of the murine coronavirus nucleocapsid protein. *J. Virol.* **83**, 7221–7234.
31. Almazan, F., Gonzalez, J. M., Penzes, Z., Izeta, A., Calvo, E., Plana-Duran, J. *et al.* (2000). Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc. Natl Acad. Sci. USA*, **97**, 5516–5521.
32. Stalcup, R. P., Baric, R. S. & Leibowitz, J. L. (1998). Genetic complementation among three panels of mouse hepatitis virus gene 1 mutants. *Virology*, **241**, 112–121.
33. Mascotti, D. P. & Lohman, T. M. (1990). Thermodynamic extent of counterion release upon binding oligolysines to single-stranded nucleic acids. *Proc. Natl Acad. Sci. USA*, **87**, 3142–3146.
34. Chen, X., Agarwal, A. & Giedroc, D. P. (1998). Structural and functional heterogeneity among the zinc fingers of human MRE-binding transcription factor-1. *Biochemistry*, **37**, 11152–11161.
35. Record, M. T., Ha, J.-H., Fisher, M. A. & Robert, T. S. (1991). Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods Enzymol.* **208**, 291–343.
36. Saikatendu, K. S., Joseph, J. S., Subramanian, V., Neuman, B. W., Buchmeier, M. J., Stevens, R. C. *et al.* (2007). Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-terminal domain of N protein. *J. Virol.* **81**, 3913–3921.
37. Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P. *et al.* (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science*, **300**, 1394–1399.
38. Marra, M. A., Jones, S. J. M., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S. N. *et al.* (2003). The genome sequence of the SARS-associated coronavirus. *Science*, **300**, 1399–1404.
39. Yount, B., Roberts, R. S., Lindesmith, L. & Baric, R. S. (2006). Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: Engineering a recombination-resistant genome. *Proc. Natl Acad. Sci. USA*, **103**, 12546–12551.
40. Liu, P., Li, L., Millership, J. J., Kang, H., Leibowitz, J. L. & Giedroc, D. P. (2007). A U-turn motif-containing stem-loop in the coronavirus 5′ untranslated region plays a functional role in replication. *RNA*, **13**, 763–780.
41. Myong, S., Cui, S., Cornish, P. V., Kirchhofer, A., Gack, M. U., Jung, J. U. *et al.* (2009). Cytosolic Viral Sensor RIG-I is a 5′-triphosphate-dependent translocase on double-stranded RNA. *Science*, **323**, 1070–1074.
42. Auweter, S. D., Oberstrass, F. C. & Allain, F. H.-T. (2006). Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Res.* **34**, 4943–4959.
43. Spriggs, S., Garyu, L., Connor, R. & Summers, M. F. (2008). Potential intra-and intermolecular interactions involving the unique-5′ region of the HIV-1 5′-UTR. *Biochemistry*, **47**, 13064–13073.
44. Tan, Y. W., Fang, S., Fan, H., Lescar, J. & Liu, D. X. (2006). Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells. *Nucleic Acids Res.* **34**, 4816–4825.
45. Narayanan, K., Kim, K. H. & Makino, S. (2003). Characterization of N protein self-association in coronavirus ribonucleoprotein complexes. *Virus Res.* **98**, 131–140.
46. Hsieh, P.-K., Chang, S. C., Huang, C.-C., Lee, T.-T., Hsiao, C.-W., Kou, Y.-H. *et al.* (2005). Assembly of Severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. *J. Virol.* **79**, 13848–13855.
47. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. (2001). A mechanism for initiating RNA-dependent RNA polymerization. *Nature*, **410**, 235–240.
48. Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F. & Weber, P. C. (1999). Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat. Struct. Mol. Biol.* **6**, 937–943.
49. Chinnaswamy, S., Yarbrough, I., Palaninathan, S., Kumar, C. T. R., Vijayaraghavan, V., Demeler, B. *et al.* (2008). A locking mechanism regulates RNA synthesis and host protein interaction by the hepatitis C virus polymerase. *J. Biol. Chem.* **283**, 20535–20546.
50. Schelle, B., Karl, N., Ludewig, B., Siddell, S. G. & Thiel, V. (2005). Selective replication of coronavirus genomes that express nucleocapsid protein. *J. Virol.* **79**, 6620–6630.
51. van Hemert, M. J., van den Worm, S. H. E., Knoops, K. v., Mommaas, A. M., Gorbalenya, A. E. & Snijder, E. J. (2008). SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity *in vitro*. *PLoS Pathog.* **e1000054**, 4.
52. Nixon, P. L., Rangan, A., Kim, Y. G., Rich, A., Hoffman, D. W., Hennig, M. *et al.* (2002). Solution structure of a luteoviral P1-P2 frameshifting mRNA pseudoknot. *J. Mol. Biol.* **322**, 621–633.
53. Sheffield, P., Garrard, S. & Derewenda, Z. (1999). Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Prot. Express Purif.* **15**, 34–39.
54. VanZile, M. L., Cosper, N. J., Scott, R. A. & Giedroc, D. P. (2000). The zinc metalloregulatory protein *Synechococcus* PCC7942 SmtB binds a single zinc ion per monomer with high affinity in a tetrahedral coordination geometry. *Biochemistry*, **39**, 11818–11829.
55. Painter, J. & Merritt, E. A. (2006). TLSMD web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.* **39**, 109–111.
56. Kuzmic, P. (1996). Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260–273.
57. MicroCal. (2002). MicroCalorimeter User's Manual Northampton, MA.
58. Yount, B., Denison, M. R., Weiss, S. R. & Baric, R. S. (2002). Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J. Virol.* **76**, 11065–11078.
59. Johnson, R. F., Feng, M., Liu, P., Millership, J. J., Yount, B., Baric, R. S. *et al.* (2005). The effect of mutations in the mouse hepatitis virus 3′(+)42 protein binding element on RNA replication. *J. Virol.* **79**, 14570–14585.