JOURNAL OF VIROLOGY, June 2007, p. 6700–6708 Vol. 81, No. 12
0022-538X/07/$08.00+0 doi:10.1128/JVI.02817-06
Copyright © 2007, American Society for Microbiology. All Rights Reserved.

Crystal Structure of a Monomeric Form of Severe Acute Respiratory Syndrome Coronavirus Endonuclease nsp15 Suggests a Role for Hexamerization as an Allosteric Switch▽

Jeremiah S. Joseph,[1]† Kumar Singh Saikatendu,[1]† Vanitha Subramanian,[1] Benjamin W. Neuman,[2] Michael J. Buchmeier,[2] Raymond C. Stevens,[3] and Peter Kuhn[1]*

Department of Cell Biology,[1] Molecular and Integrative Neurosciences Department,[2] and Department of Molecular Biology,[3] The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037

Received 20 December 2006/Accepted 26 March 2007

Mature nonstructural protein-15 (nsp15) from the severe acute respiratory syndrome coronavirus (SARS-CoV) contains a novel uridylate-specific Mn^{2+}-dependent endoribonuclease (NendoU). Structure studies of the full-length form of the obligate hexameric enzyme from two CoVs, SARS-CoV and murine hepatitis virus, and its monomeric homologue, XendoU from *Xenopus laevis*, combined with mutagenesis studies have implicated several residues in enzymatic activity and the N-terminal domain as the major determinant of hexamerization. However, the tight link between hexamerization and enzyme activity in NendoUs has remained an enigma. Here, we report the structure of a trimmed, monomeric form of SARS-CoV nsp15 (residues 28 to 335) determined to a resolution of 2.9 Å. The catalytic loop (residues 234 to 249) with its two reactive histidines (His 234 and His 249) is dramatically flipped by ~120° into the active site cleft. Furthermore, the catalytic nucleophile Lys 289 points in a diametrically opposite direction, a consequence of an outward displacement of the supporting loop (residues 276 to 295). In the full-length hexameric forms, these two loops are packed against each other and are stabilized by intimate intersubunit interactions. Our results support the hypothesis that absence of an adjacent monomer due to deletion of the hexamerization domain is the most likely cause for disruption of the active site, offering a structural basis for why only the hexameric form of this enzyme is active.

Nidoviruses are enveloped, positive-stranded RNA viruses consisting of the families *Coronaviridae*, *Arteriviridae*, and *Roniviridae* (5, 34). These viruses have large genomes, and their gene expression is controlled by a complex and poorly understood membrane-anchored replicase/transcriptase complex (12, 33). Components of this complex originate from the large 5′-proximal open reading frame 1a (ORF1a) and ORF1b, which span two-thirds of the genome. ORF1a is translated into polyprotein 1a (pp1a), while polyprotein 1ab (pp1ab) is formed when −1 ribosomal frameshifting just upstream of the ORF1a stop codon causes read-through into ORF1b. In the severe acute respiratory syndrome coronavirus (SARS-CoV), these polyproteins are proteolytically processed by two virus-encoded proteases (a main protease and a papain-like protease) into 16 mature nonstructural proteins (nsp1 to -16) (9, 23, 29). These proteins act in concert to replicate the viral genome and transcribe a nested set of eight subgenomic mRNAs, which are then translated to produce the structural and accessory proteins. One of the RNA-processing enzymes in the viral replicase-transcriptase is a uridylate-specific endoribonuclease (NendoU), which is considered a genetic marker for the nidoviruses, discriminating them from all other RNA virus families (14). This protein is a distant homologue of an endoribonuclease, XendoU from *Xenopus laevis*, that is involved in processing intron-encoded box C/D U16 small nucleolar RNA (10, 17) and shares most of its catalytic determinants.

nsp15 of SARS-CoV is a 346-residue polypeptide that results from the cleavage of pp1ab at sites ^{6427}RLQ↓SLE^{6432} and ^{6773}KLQ↓ASQ^{6778} by the main protease (nsp5). Several recent studies have focused on structural and functional characterization of coronavirus nsp15 due to its potential importance as a drug target. Mutagenic inactivation of this enzyme renders the virus nonviable, as demonstrated with human CoV 229E (14) and equine arteritis virus (EAV; reference 12). nsp15 preferentially cleaves the 3′ end of uridylates of RNA at GUU or GU sequences to produce molecules with 2′-3′ cyclic phosphate ends (2, 14). It acts on both double-stranded RNA and single-stranded RNA (ssRNA) but with different preferences (2, 14). Mn^{2+} ions are required for optimal activity (2, 14), with the ion apparently binding weakly to the protein but producing significant conformational changes upon binding (2). The recent structures of full-length nsp15 from SARS-CoV (28) and murine hepatitis virus JHM strain (MHV-JHM) (38) and its eukaryotic homolog, XendoU from *Xenopus laevis* (27), have provided first structural and mechanistic descriptions of this new enzyme family. Its catalytic center resembles the active site of an unrelated nuclease, RNase A (28).

The biological unit of CoV nsp15 is a hexamer (13) (see Fig. 3a), with its six (potential) active sites distributed along its periphery away from intersubunit interfaces. nsp15 exists in solution in an equilibrium between monomers and hexamers (and sometimes other oligomeric states), but only the hexameric form has been reported to be enzymatically active (13, 38). The link between its hexameric state and its enzymatic

* Corresponding author. Mailing address: 10550 N. Torrey Pines Road, CB265, The Scripps Research Institute, La Jolla, CA 92037. Phone: (858) 784-9114. Fax: (858) 784-8996. E-mail: pkuhn@scripps.edu.

† J.S.J. and K.S.S. contributed equally to this work.

▽ Published ahead of print on 4 April 2007.

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

activity has not been clear. However, based on our structure of a trimmed monomeric form of SARS-CoV nsp15 (lacking 28 N-terminal and 11 C-terminal residues), we demonstrate evidence for the structural basis underlying this. Examination of the active site strongly suggests that the absence of monomer-monomer interactions within the hexamer destabilizes and significantly displaces two loops (residues 234 to 249 and 276 to 295) in the catalytic domain, resulting in the destruction of the active site in the isolated monomer.

MATERIALS AND METHODS

Construct design and cloning. The sequence encoding the 346-residue nsp15 (NP_828872.1; gi:29837507) corresponds to nucleotides 19551 to 20588 in the genome of the SARS-CoV Tor2 strain (nomenclature of nsp's is as per the work of Snijder and coworkers [33]). A construct corresponding to residues 28 to 346 of nsp15 was amplified by PCR from genomic cDNA of the SARS-CoV Tor2 strain by use of *Taq* polymerase and primer pairs containing the predicted 5′ and 3′ ends (forward, 5′-ATGAATAATGCTGTTTACACAAAGGTA GATGGTA GGGCCGGCCGGG-3′; reverse, 5′-TTGTAGTTTTGGGTAGAAGGTTT CA ACATGTCCCCGGCCGGCCCTA-3′). The PCR product was cloned between FseI and PmlI sites into the expression vector pMH1F, a derivative of pBAD (Invitrogen). Expression in pMH1F is driven by the araBAD promoter, and the recombinant protein is produced with a short, noncleavable N-terminal $Thio_6His_6$ tag (MGSDKIHHHHHH). As part of the crystallization and diffraction optimization strategy, C-terminal truncation mutants were generated by insertion of appropriate stop codons into the coding sequence by site-directed mutagenesis using a Stratagene QuikChange kit. The construct described in this paper corresponds to residues 28 to 335.

Expression and purification. A sequence-verified clone was transformed into Top10 cells (Invitrogen). An overnight culture from a fresh transformant was used to inoculate flasks of 2XYT-ampicillin medium. The culture was grown at 37°C to an optical density at 600 nm of 0.6, induced with 0.2% (wt/vol) L-arabinose, and further grown at 14°C overnight. The cells were harvested by centrifugation and lysed by sonication in buffer containing 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% glycerol, 0.5 mg/ml lysozyme, 100 μl/liter benzonase, and EDTA-free protease inhibitor (one tablet per 50 ml buffer; Roche). The lysate was clarified by ultracentrifugation at 45,000 rpm for 20 min at 4°C, and the soluble fraction was applied onto a metal chelate column (Talon resin charged with cobalt; Clontech). The column was washed with 20 mM Tris-HCl, pH 7.8, 300 mM NaCl, 10% glycerol, 5 mM imidazole and eluted with 25 mM Tris-HCl, pH 7.8, 300 mM NaCl, 150 mM imidazole. The eluate was then fractionated by anion exchange on a Poros HQ column using a linear gradient of NaCl (0 to 1 M) in 25 mM Tris-HCl, pH 8.0. The pure fractions were pooled and concentrated to 1.8 mM. The protein was either flash frozen in liquid nitrogen for later use or used immediately for crystallization trials.

Crystallization and data collection. Crystals were grown by the nanovolume sitting drop method (30) using drops consisting of 100 nl 1.8 mM protein and 100 nl crystallant. Large (~300-μm by ~20-μm by ~200-μm) crystals of hexagonal morphology grew typically overnight in 0.2 M sodium bromide, 0.1 M sodium acetate, pH 5.5, and 25% polyethylene glycol (PEG) 2000MME. These were cryoprotected in a solution containing mother liquor and 15% glycerol and flash frozen in liquid nitrogen. While optimizing diffraction, crystals were screened at the Stanford Synchrotron Radiation Laboratory beamlines 1-5, 11-1, 11-3, and 9-1 and the General Medicine and Cancer Institutes Collaborative Access Team (GM/CA-CAT) beamline 23-ID of the Advanced Photon Source (Argonne, IL) by use of Blu-Ice (19). The complete 2.9-Å native data set used for phasing and refinement was collected at the GM/CA-CAT. The reflections were indexed in the primitive hexagonal lattice scaled in P3 using HKL2000 (21). Subsequent molecular replacement attempts (see below) revealed the actual space group to be $P3^1$.

Phasing and refinement. Initial phases were obtained by molecular replacement using the full atom monomer structure of full-length SARS-CoV nsp15 (kindly provided by Bruno Canard) with the program Phaser (26) within CCP4 (6) using data from 50.0 to 2.9 Å. Rigid-body refinement using Refmac5 (20) revealed a clearly interpretable electron density map. Phases and the model itself were further improved by one round of Arp/wARP (16). Initial model building was guided by a composite omit map calculated using CNS (4) to minimize model bias. Subsequent improvement of the coordinates was achieved by manual model building in Coot (8) alternating with restrained refinement using Refmac5. Though noncrystallographic symmetry was identified between the four monomers, they were not used during refinement due to subtle differences observed between them in the asymmetric unit. The final model statistics, validation, and stereochemical quality are summarized in Table 1.

TABLE 1. Data collection, magnetic resonance, and refinement statistics for nsp15

Parameter[a]	Finding or value[b]
Data collection	
Lattice	Primitive hexagonal
Space group	$P3^1$
Unit cell	
α (90°)	98.974 Å
β (90°)	98.974 Å
γ (120°)	214.926 Å
Wavelength	0.9793 Å
Resolution range	85–2.9 (3.0–2.90)
Total observations	581,802
Unique reflections	52,135 (3,114)
Completeness	98.1% (89.9%)
Redundancy	2.9 (1.9)
Mean I/σ value	14.4 (1.44)
R_{sym} on I value	0.085 (0.466)
Refinement statistic	
Resolution range	50–2.90 Å
R_{cryst} value	0.251 (0.341)
R_{free} value	0.292 (0.429)
Bond length RMSD	0.015 Å
Bond angle RMSD	1.98°
Average isotropic B value	43.1 $Å^2$

[a] RMSD, root mean square deviation. Equations: $R_{sym} = \Sigma hkl[(\Sigma j\, Ij - \langle I \rangle)/\Sigma j\, Ij]$; $R_{work} = \Sigma hkl\, Fo - Fc / \Sigma hkl\, Fo$, where Fo and Fc are the observed and calculated structure factors. Five percent of randomly chosen reflections were used to calculate R_{free}.

[b] Values in parentheses are for data corresponding to the outermost reflection shell.

Electrophoretic mobility shift assay. Electrophoretic mobility shift assays were used to assess RNA binding by the monomeric nsp15 construct. A short, 5′-fluoresceinated RNA oligomer (5′-GGUUU-3′) was designed to mimic the preferred recognition site of nsp15 (14). The positive control, purified SARS-CoV nsp10, was produced as described previously (15). Protein samples were mixed with 0.8 μg of RNA in assay buffer containing 150 mM NaCl, 50 mM Tris-HCl, and 5 mM $CaCl_2$, pH 8.0. Samples were incubated at 37°C for 1 h and analyzed by native electrophoresis on 6% acrylamide DNA retardation gels (Invitrogen). Fluoresceinated RNA was visualized using a UV light source equipped with a digital camera. Protein was detected by SYPRO-ruby poststain according to the manufacturer's protocol (Molecular Probes). Densitometric analysis was performed using a flatbed scanner with ImageJ software (NIH). The amount of bound protein was calculated relative to the maximum binding observed in each experiment. K_d values were determined from the midpoints of the fitted titration data as described elsewhere (32).

PFO-PAGE. Perfluoro-octanoic acid polyacrylamide gel electrophoresis (PFO-PAGE) was performed according to the method of Ramjeesingh et al. (25) to assess protein stoichiometry. Purified protein samples were mixed 1:1 with PFO loading buffer containing 8% (wt/vol) PFO, 100 mM Tris base, 20% (vol/vol) glycerol, and 0.05% (wt/vol) orange G. Samples were loaded onto precast 4 to 20% Tris-glycine gels, and electrophoresis was performed with a standard Tris-glycine running buffer to which 0.5% (wt/vol) PFO was added. Gels were stained with SYPRO-ruby according to the manufacturer's instructions and photographed with a digital camera attached to a UV light source.

Protein structure accession number. The structure factors and coordinates of the structure have been deposited in the Protein Data Bank (PDB) with accession number 2OZK.

RESULTS AND DISCUSSION

Expression and crystallization of monomeric nsp15 construct. We obtained high expression (final yield, 35 mg/liter) of a pMH1F construct corresponding to nsp15 residues 28 to 346

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

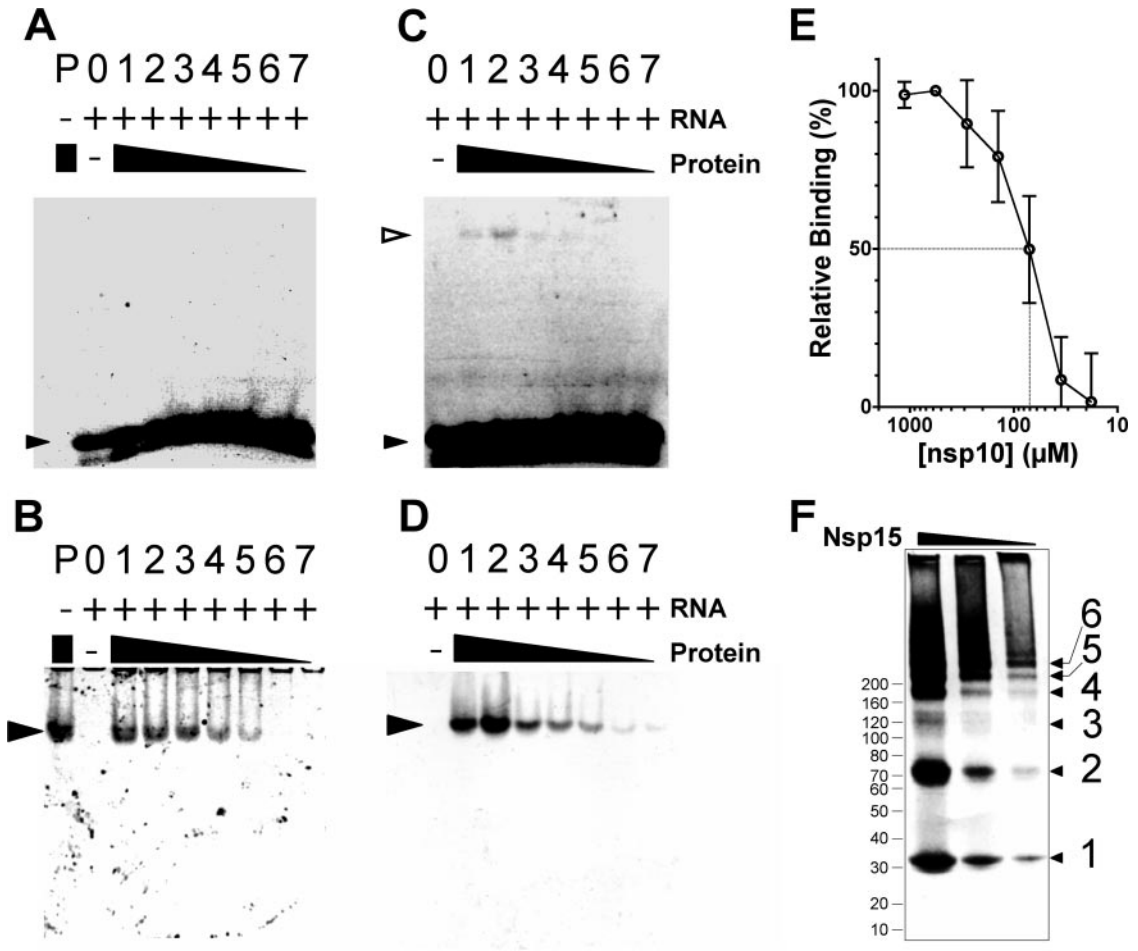

FIG. 1. Electrophoretic mobility shift assays. Monomeric SARS-CoV nsp15 (A and B) and native nsp10 (C and D) were incubated with fluoresceinated ssRNA substrate, and complexes were separated by native electrophoresis on 6% polyacrylamide gels. Fluoresceinated RNA was visualized under UV illumination (A and C), and protein was visualized similarly after SYPRO-ruby poststaining (B and D). Lanes: P, protein only; 0, RNA only; 1 to 7, gradient of protein with a fixed amount of RNA. Free RNA is indicated with a small filled triangle, protein is indicated with a large filled triangle, and protein-RNA complexes are indicated with open triangles. (E) nsp10 was used as a positive control; its binding affinity to the fluoresceinated RNA substrate is shown. The dotted lines indicate the concentration of nsp10 that correlates to 50% binding. (F) PFO-PAGE gel indicating that trimmed nsp15 exists primarily as a monomer and as a dimer; although a variety of larger forms exists, this construct appears to be specifically defective in trimer formation.

in Top10 cells. This construct lacks the first 28 residues (in particular Glu 3) of the N-terminal domain, which are important for oligomerization (13). It eluted predominantly as a monomer in gel filtration studies (data not shown), although small amounts of higher oligomeric forms were detected after PFO-PAGE analysis (Fig. 1F).

This construct crystallized overnight in several PEG-containing conditions but diffracted poorly (~6 Å). We expressed several C-terminal truncation mutants, deleting from 5 to 19 residues, all of which had comparably high expression levels and were monomeric in solution. Among these, a construct from residues 28 to 335 (described in the rest of this paper) crystallized in a non-PEG condition, which upon optimization with additive screens yielded crystals that diffracted to 2.9 Å. Attempts to improve diffraction using annealing or dehydration (including the procedure described for MHV nsp15 in reference 38) were not successful.

It is interesting that while our enzymatically inactive (see below) truncation mutant of nsp15 expressed in high amounts, the full-length wild-type (WT) construct expressed only weakly (data not shown). Ricagno et al. also reported low expression for WT SARS-CoV nsp15 (28). These results taken together reinforces the suggestion by Xu et al. that expression of WT MHV-JHM nsp15 may be toxic to *Escherichia coli*, causing slow cell growth and low protein yields (38). An incidental mutation, F307L, and active site mutations H262S, H277S, and K317S—all detrimental to enzymatic activity—expressed at much higher levels. Even in EAV NendoU, only the inactive D3014A mutant could be expressed in *E. coli*, whereas WT expression was not tolerated (22). It is likely that functional nsp15 may act on its own (and also cellular) mRNA, limiting its expression and impeding normal cell growth.

Functional characterization of monomeric nsp15. Bhardwaj et al. reported that only the hexameric form of full-length WT nsp15 binds RNA and that Mn^{2+} ions increase the affinity of interaction (1). We therefore assessed the ability of mono-

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

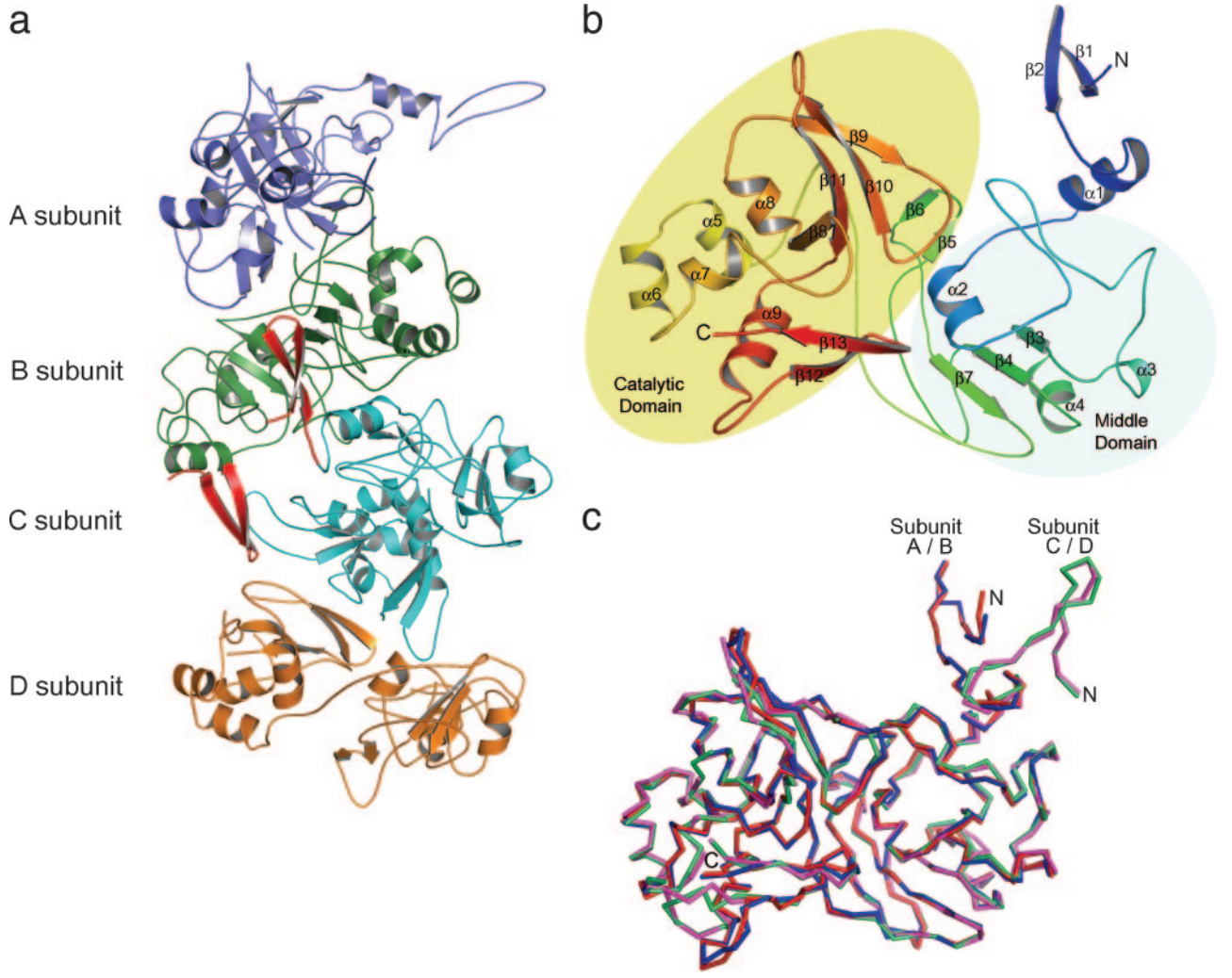

FIG. 2. (a) Contents of the asymmetric unit. The four monomers (designated A, B, C, and D) are colored blue, green, cyan, and orange, respectively. (b) The three-domain topology of the truncated SARS-CoV nsp15 monomer. Color is ramped from the N terminus (blue) to the C terminus (red). Secondary structural elements are numbered β1 to β13 for β-strands, while α-helices are numbered α1 to α9. (c) Superimposition of the four monomers of the asymmetric unit. Subunits A/B, which form a pair, are colored blue and red, while the other two subunits, C and D, are colored green and magenta, respectively.

meric nsp15 construct 28-346 and of seven additional truncation constructs (lacking 5, 7, 9, 11, 13, 15, 17, and 19 C-terminal residues, respectively) to bind and cleave ssRNA. Electrophoretic mobility shift assays did not reveal an interaction between any form of monomeric nsp15 and fluoresceinated ssRNA substrates (Fig. 1), compatible with previous reports (1, 2). These results confirmed that monomeric nsp15 lacks RNA binding activity, as previously reported (1, 2). We do not observe any effect of Mn^{2+} ions on the affinity of monomeric nsp15 for RNA (data not shown). The modulation of RNA binding by Mn^{2+} ions is therefore probably a feature only of the intact nsp15 hexamer. XendoU, a structural and functional homologue of viral NendoU from *Xenopus laevis*, is biologically active as a monomer and does bind RNA in the absence of Mn^{2+} ions (10). However, changes to the affinity of this interaction in the presence of Mn^{2+} ions have not been investigated.

Description of the asymmetric unit. The structure of SARS-CoV nsp15 reported here covers residues 28 to 335 of nsp15 from the Tor2 strain (Fig. 2). Data collection and refinement statistics are summarized in Table 1. Cell content analysis yields a Matthews coefficient of 5.06 and a large solvent content of 74%, which might explain the poor diffraction of these otherwise large (0.3-mm) crystals. The final refined model contains four monomers arranged coaxially in the asymmetric unit of the hexagonal $P3^1$ lattice (Fig. 2a). The four monomers exist as two pairs, A/B and C/D. Monomers A and B are virtually identical to each other but differ from monomers C and D (Fig. 2c). The N terminal β-hairpin is tilted by ~45° in monomers A/B in comparison to the pair C/D. None of the intersubunit interactions seen between monomers within the asymmetric unit as well as with their symmetry mates recapitulate those observed for the functionally relevant hexamer of full-length nsp15 of SARS-CoV (Fig. 3a) or MHV.

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

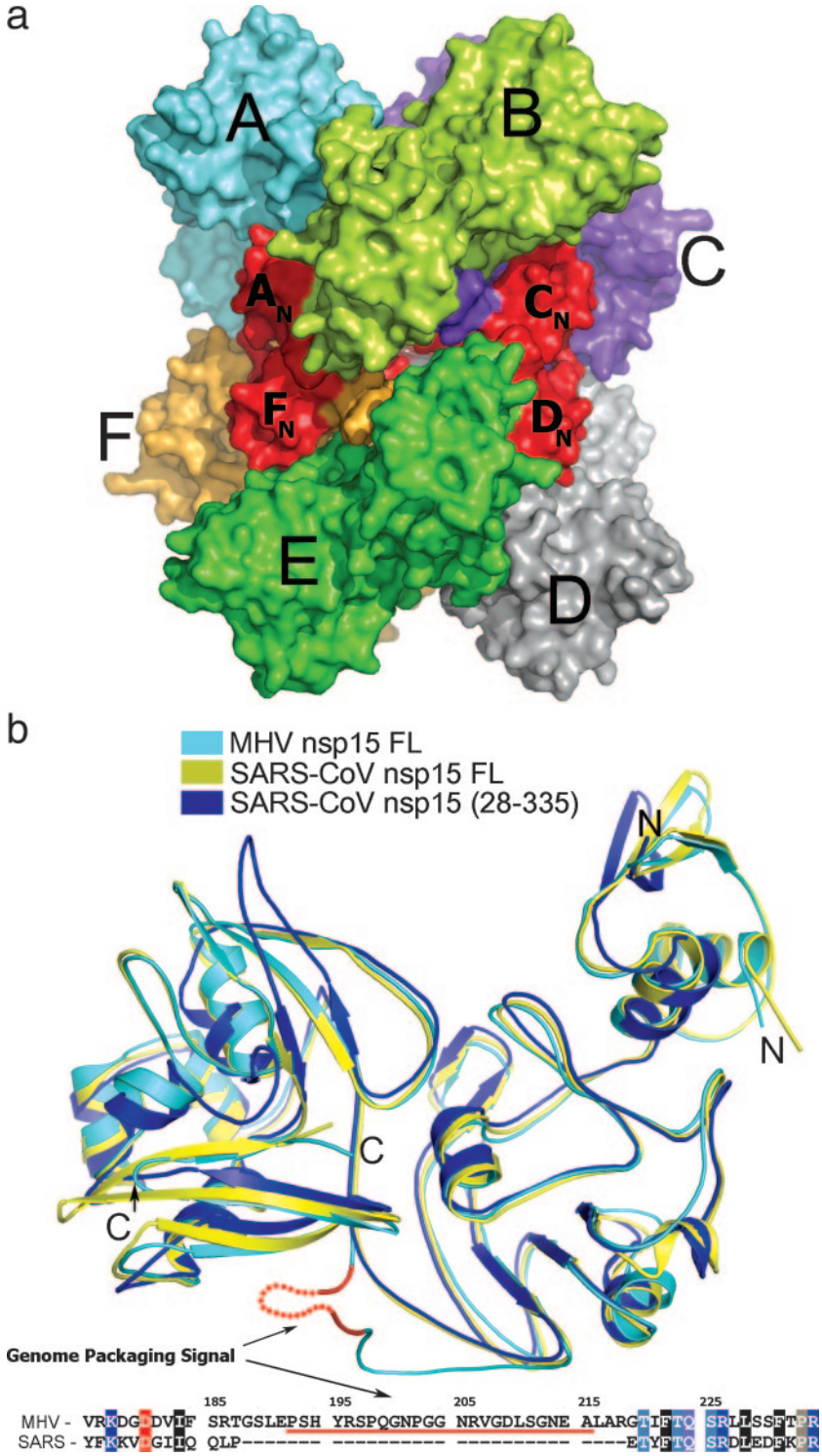

FIG. 3. (a) Surface representation of hexameric full-length (FL) nsp15 (PDB code 2H85). The individual subunits are colored and marked A to F. The N-terminal 28 residues of the four monomers A, C, D, and F are shown in red and labeled A_N, C_N, D_N, and F_N, respectively. (b) Superposition of truncated (blue) and full-length (yellow) SARS-CoV nsp15 with MHV nsp15 (cyan). The genomic packaging signal of MHV is colored red, and the disordered region is shown as a dotted line. The sequence alignment of the packaging signal of MHV is shown below the ribbon diagram, with the implicated region (P192 - A215) indicated with a red line. Sequence IDs are gi:29837507 for SARS-CoV and gi:37999877 for MHV. (c) Conformational rearrangement of the active site loop and supporting loop in the truncated and full-length forms of nsp15. Truncated monomeric nsp15 is shown in green, while that of the hexameric WT is in light brown. The "supporting loop" (residues 276 to 295) and the "active site loop" (residues 234 to 249) are colored red in monomeric nsp15 and blue in full-length nsp15. Equivalent residues in the active sites of the two structures are labeled (regular font in monomeric nsp15 and in italics for the full-length enzyme). (d) Electron density map (2Fo-Fc) contoured at 1.1 σ around key residues in the active site of monomeric nsp15.

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

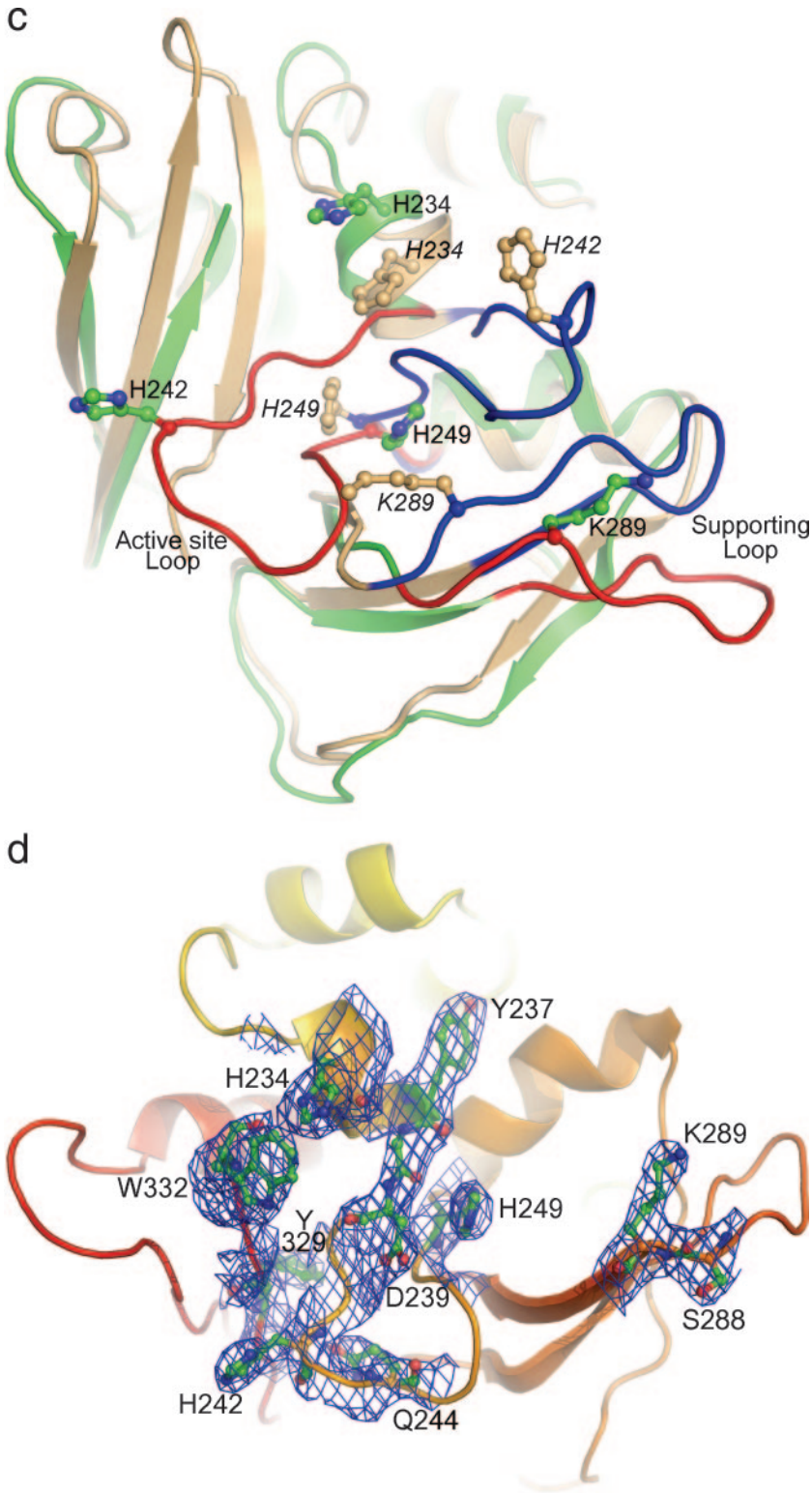

FIG. 3—*Continued*.

As anticipated, each monomer adopts an overall topology similar to that of the full-length WT nsp15, with three distinct structural domains (Fig. 2b). The first domain is incomplete in our construct and consists of a β-hairpin, which is displaced by about 90° in monomers C/D with respect to A/B (Fig. 2c). A surface representation of the full-length hexameric SARS-CoV nsp15 is shown in Fig. 3a, highlighting the role of the N terminal residues (A_N to D_N) in stabilizing the hexamer. A struc-

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

ture superposition of MHV nsp15 and the two forms of SARS-CoV nsp15 is shown in Fig. 3b. The central domain of truncated nsp15 superimposes nearly perfectly on the two full-length structures, but there are several important differences in the catalytic domain. A superposition of one of the truncated monomers on the structure of the full-length nsp15 is shown in Fig. 3c. Polypeptide segments corresponding to residues 240 to 247 and 282 to 286 are disordered only in the C and D monomers and are clearly ordered in A/B. In monomers A/B, the loop consisting of residues 308 to 317 is bent outwards in a direction away from the active site by about 70° with respect to the corresponding region in C/D. The importance of these differences is discussed later.

Structural basis for the functional hexamer. Most striking of all is the comparison between structures of full-length SARS-CoV nsp15 and our trimmed construct, particularly in the catalytic subunit (Fig. 3c). In monomers A (and B) in the asymmetric unit of our construct, the loop consisting of residues 234 to 249 (referred to here as the "active site loop") spanning the two active site histidines (His 234 and His 249) is flipped by ~120° into the active site cleft. Lys 289, another active site residue, points in a diametrically opposite direction (a consequence of a significant displacement of the loop from residues 276 to 295), making it very unlikely to be available for catalysis. Furthermore, it is precisely in each of these regions that significant differences were observed between monomers A/B and C/D in the asymmetric unit of our crystal structure. In C (and D), electron density for these loops is missing, indicating that they are flexibly disordered. In the structures of the functional full-length nsp15 hexamer, the "active site loop" and the "supporting loop" are packed against each other and are stabilized by intimate interactions with residues in the adjacent monomer. Although our structure is of a nominal (2.9-Å) resolution, the electron density for these two loops is quite well ordered in two of the monomers and clearly traceable after residues were deleted in early stages of model improvement as well as in the composite omit map. The electron density around the active site residues is shown in Fig. 3d. Residues Asp239 - Gly246 and Leu248 in the active site loop interact with Ile280, Asp 282, Thr 285 - Lys289, Ser293, and Val294 in the "supporting loop." The supporting loop, in turn, makes extensive interactions with the adjacent monomer (for example monomer A with monomer E in the hexamer; PDB 2H85). Specifically, Lys264, Glu266, Phe268 - Met271, Asn277, Phe279 - Ser288, Cys290, and Gly291 from the supporting loop interact with five sequentially distinct regions on the adjacent monomer (Val9 - His14, Asp16, Thr33, Val35, Val40 - Ile42, Arg61, Ile63, Tyr88, Val162 - Val165, Leu167 - Val172), which form a contiguous patch at the intersubunit interface. In the structure of the truncated form described here, the absence of the adjacent monomer likely causes a peeling away of the "supporting loop" from residues 276 to 295, causing an apparent inward collapse of the "active site loop" from residues 234 to 249 into the active site, thus possibly destroying it. Alternatively, just disorder in this region, as found in chains C and D in our crystal structure, may contribute to a destruction of a productive active site conformation. In chains C and D, the "active site loop" and the "supporting loop" face the large solvent channels in the crystal, while they are stabilized in their "inactive" conformation by crystal contacts in chains A and B.

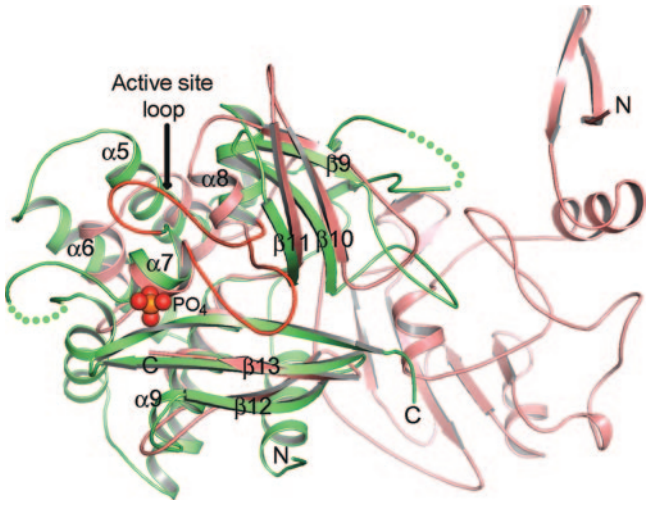

FIG. 4. Superposition of XendoU structure (PDB ID 2C1W; green) onto the catalytic domain of SARS-CoV NendoU (PDB ID 2H85; light pink). The two disordered loop regions in the XendoU structure are indicated by dotted lines. The secondary structures are numbered using the same scheme as for Fig. 2b. The active site loops of the two structures are highlighted in red.

The eukaryotic structural homologue of NendoU from *Xenopus laevis* (referred to as XendoU in many studies) exists as a monomer and yet is enzymatically active (10). Structural comparison reveals that the single domain of XendoU corresponds to the catalytic domain of CoV nsp15. Most secondary structure elements of SARS-CoV nsp15 superimpose well with XendoU (Fig. 4). However, there are several additional structural elements in XendoU which decorate the periphery of the molecule, for instance, the helix comprised of residues 10 to 21 followed by a long hairpin loop held together by a short two-stranded β-sheet (strand 1, residues 32 to 35; strand 2, residues 60 to 62) and a short helical turn (residues 64 to 67). Two helices in the SARS-CoV structure (residues 206 to 213 and 219 to 225) are extended in XendoU (residues 73 to 86 and 96 to 111, respectively). These extensions with a longer connecting loop (10 residues in XendoU compared to 6 residues in SARS-CoV NendoU) act as one of the walls of the monomer, providing stability to the active site, possibly by preventing it from splaying open or collapsing. Two additional helices from residues 113 to 125 and from 133 to 144 provide additional undergirding to the molecule and play a role in stabilizing the active site loop by holding up the loop from residue 165 to 178. Further, residues 145 to 158 form a hairpin loop that also provides stability, holding up the loop consisting of residues 165 to 178. These additional structural elements hold the cup-like molecule together, likely fulfilling the role of adjacent monomers in the CoV nsp15 hexamer. Hence, a comparison of our structure of the monomeric form of nsp15 with those of full-length CoV nsp15 and monomeric XendoU offers a likely structural basis for the catalytic activity of the intact CoV nsp15 hexamer only.

It is fair to assume that nsp15 exists as a monomer (perhaps as part of a long polyprotein precursor) before it assembles as a hexamer. One would therefore expect that in its polyprotein precursor, or when freshly released from it by the proteolytic

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

action of nsp5, nsp15 would bear the same conformationally inactive catalytic site—the default "off" position—in the absence of the other "fortifying" monomers of the hexamer. Translation of additional monomers and subsequent hexamerization would then switch the protein to the "on" position. In vitro, nsp15 does occur in equilibrium between the hexamer form and lower oligomeric states (13); hence, disruption of the hexamer would be a conceivable means of regulating the effective concentration of the enzyme, though there is not yet experimental evidence for this. It is worth noting that the pp1b region containing nsp12 to -16 is accessed by only up to 30% of translating ribosomes (3). Since nsp15 is absolutely required for viral RNA synthesis (13, 22), a minimum of 18 translation events may be required before the phase of RNA synthesis involving nsp15 can effectively occur. Therefore, one would anticipate that a high level of protein synthesis is a prerequisite for the discontinuous phase of coronaviral RNA synthesis (31). This in turn raises the interesting implication that CoVs may be unusually dependent on translation compared to other viruses. Whether this is indeed the case, however, remains to be experimentally ascertained.

Structural basis for inhibitor design. The activity of nsp15 is critical for the viability of the virus, making it potentially an excellent drug target. As mentioned earlier, mutation of catalytic residues of the NendoUs in human CoV 229E (14) and EAV (22) inactivates the viruses. As described earlier, the monomeric form of nsp15 is inactive due to the loop consisting of residues 234 to 249 falling into the active site. Since this loop fits well into the active site, designed peptide sequences derived from this segment could also be reasonably expected to bind to the active site and could provide a good starting point for the design of inhibitory peptidomimetic molecules, an avenue that is being actively pursued by our group. A second avenue for inhibitor design that is being explored is based on peptides or small molecules that inhibit oligomerization by disruption of the interface. These could be more desirable than active site inhibitors, which could potentially interfere with the similar active sites (such as that of RNase A) in the host cell. They would merely disrupt protein-protein interactions essential for nsp15 hexamerization and hence the activity of the enzyme.

The MHV genomic RNA packaging signal is embedded in the nsp15 ORF. A striking difference between the MHV and SARS-CoV nsp15 structures occurs in the region linking the catalytic and central domains (Fig. 3b). The MHV nsp15 interdomain linker is long and relatively flexible, as evidenced by the lack of observable electron density in the structure, in contrast to the short, structured SARS-CoV linker (28, 38). The MHV genomic RNA packaging signal, which functions at the level of RNA sequence and secondary structure, maps entirely within the extended MHV nsp15 interdomain linker (24). CoVs belonging to group 2a, which includes MHV, are all predicted to encode long nsp15 interdomain linkers with elevated proline and glycine content (25 to 33% P+G), and the genomic RNA packaging signal of bovine CoV maps to this region (7). In contrast, CoVs of groups 1, 2b, and 3 are predicted to contain shorter nsp15 interdomain linkers of the type found in SARS-CoV. Phylogenetic and structural evidence therefore suggests that the nsp15 ORF contains both a likely RNA packaging signal and a likely flexible interdomain protein loop in the group 2a CoVs only.

Several instances of signals that function at the level of RNA structure but are encoded within ORFs are known. Examples for which the structure of the encoded protein is not known include the hepatitis delta virus packaging signal in the large delta antigen ORF (36), the tobacco mosaic virus packaging signal in the GP4 movement protein ORF (35), and the poliovirus *cis*-acting replication element within the 2C coding region (11). The protein regions encoded by these RNA elements are characterized by elevated proline-plus-glycine (19% P+G) and aspartic acid-plus-glutamic acid (16% D+E) content, consistent with local flexibility at the level of protein structure.

Examples of long RNA signals within ORFs for which the structure of the encoded protein is known include the human rhinovirus *cis*-acting replication element (reference 18; PDB code 4RHV), black beetle virus and Nodamura virus packaging signals (reference 39; PDB codes 2BBV and 1NOV, respectively), and the MHV packaging signal (reference 37; PDB code 2GTH). These regions are also collectively enriched in proline and glycine (23% P+G) and adopt mainly randomly coiled conformations. As would be expected from the amino acid content, only 16% of the amino acid residues encoded by these RNA elements participate in α-helical and β-strand secondary structures in the crystallized proteins. Therefore, from the limited data available it would appear that long RNA signals located within protein-encoding regions generally encode randomly coiled, proline- and glycine-rich protein loops.

CONCLUSION

Here we have presented the structure of a truncated, monomeric form of SARS-CoV nsp15 that lacks the N terminal hexamerization domain. The resulting monomerization led to the collapse of the active site in the catalytic domain. This has provided a mechanistic explanation for the critical role of hexamerization in the enzymatic activity of the full-length coronaviral nsp15. This study also presents a structural basis for the possible design of novel inhibitors that target the active site as well as interfere with hexamerization. Finally, we point out that the RNA packaging signals in group 2a CoVs (and several other viruses) map to unstructured protein loops, offering a means for independent adaptation of the RNA packaging signal without materially altering the structure of the encoded protein.

ACKNOWLEDGMENTS

We thank Bruno Canard and Zihe Rao for providing prereleased coordinates of the SARS-CoV and MHV NendoU for structure analysis and MR trials. We acknowledge Tom Clayton, Mark Griffith, Kin Moy, and Anna Tang for help with protein purification; Alexei Brooun, Christine Chen, and Jeffrey Velasquez for help in cloning; and Joseph Klaus for assistance with electrophoretic mobility shift assays and PFO-PAGE. Bioinformatics support for this project was provided by Enrique Abola, Anand Kolatkar, and Sophie Coon of The Scripps Research Institute. Support provided by the staff scientists at SSRL and APS GM/CA-CAT (23-ID) during on-site as well as remote crystal screening and data collection is gratefully acknowledged.

SSRL beamline BL11-1 is supported by the National Institutes of Health (NIH) National Center for Research Resources, NIH National Institutes of General Medical Sciences, Department of Energy, Office of Biological and Environmental Research, Stanford University, and The Scripps Research Institute (TSRI). The GM/CA-CAT beamline

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY

(23-ID) is supported by the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (Y1-GM-1104). This study was supported by National Institutes of Allergy and Infectious Disease/NIH contract HHSN 266200400058C, "Functional and Structural Proteomics of the SARS-CoV," to P.K.

This is TSRI manuscript 18649.

REFERENCES

1. **Bhardwaj, K., L. Guarino, and C. C. Kao.** 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. J. Virol. **78:**12218–12224.
2. **Bhardwaj, K., J. Sun, A. Holzenburg, L. A. Guarino, and C. C. Kao.** 2006. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. J. Mol. Biol. **361:**243–256.
3. **Brierley, I., M. E. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. Brown, and S. C. Inglis.** 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J. **6:**3779–3785.
4. **Brunger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren.** 1998. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. Sect. D **54:**905–921.
5. **Cavanagh, D.** 1997. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch. Virol. **142:**629–633.
6. **Collaborative Computational Project, Number 4.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D **50:**760–763.
7. **Cologna, R., and B. G. Hogue.** 2000. Identification of a bovine coronavirus packaging signal. J. Virol. **74:**580–583.
8. **Emsley, P., and K. Cowtan.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D **60:**2126–2132.
9. **Gao, F., H. Y. Ou, L. L. Chen, W. X. Zheng, and C. T. Zhang.** 2003. Prediction of proteinase cleavage sites in polyproteins of coronaviruses and its applications in analyzing SARS-CoV genomes. FEBS Lett. **553:**451–456.
10. **Gioia, U., P. Laneve, M. Dlakic, M. Arceci, I. Bozzoni, and E. Caffarelli.** 2005. Functional characterization of XendoU, the endoribonuclease involved in small nucleolar RNA biosynthesis. J. Biol. Chem. **280:**18996–19002.
11. **Goodfellow, I. G., D. Kerrigan, and D. J. Evans.** 2003. Structure and function analysis of the poliovirus cis-acting replication element (CRE). RNA **9:**124–137.
12. **Gorbalenya, A. E.** 2001. Big nidovirus genome: when count and order of domains matter. Adv. Exp. Biol. Med. **494:**1–17.
13. **Guarino, L. A., K. Bhardwaj, W. Dong, J. Sun, A. Holzenburg, and C. Kao.** 2005. Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. J. Mol. Biol. **353:**1106–1117.
14. **Ivanov, K. A., T. Hertzig, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr.** 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. Proc. Natl. Acad. Sci. USA **101:**12694–12699.
15. **Joseph, J. S., K. S. Saikatendu, V. Subramanian, B. W. Neuman, A. Brooun, M. Griffith, K. Moy, M. K. Yadav, J. Velasquez, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2006. Crystal structure of nonstructural protein 10 from the severe acute respiratory syndrome coronavirus reveals a novel fold with two zinc-binding motifs. J. Virol. **80:**7894–7901.
16. **Lamzin, V. S., A. Perrakis, and K. S. Wilson.** 2001. The ARP/WARP suite for automated construction and refinement of protein models, p. 720–722. *In* M. G. Rossmann and E. Arnold (ed.), International tables for crystallography, vol. F. Crystallography of biological macromolecules. Kluwer Academic Publishers, Dordrecht, The Netherlands.
17. **Laneve, P., F. Altieri, M. E. Fiori, A. Scaloni, I. Bozzoni, and E. Caffarelli.** 2003. Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in Xenopus laevis. J. Biol. Chem. **278**:13026–13032.
18. **McKnight, K. L., and S. M. Lemon.** 1998. The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. RNA **4:**1569–1584.
19. **McPhillips, T. M., S. E. McPhillips, H. J. Chiu, A. E. Cohen, A. M Deacon, P. J. Ellis, E. Garman, A. Gonzalez, N. K. Sauter, R. P. Phizackerley, S. M. Soltis, and P. Kuhn.** 2002. Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. J. Synchrotron Radiat. **9:**401–406.
20. **Murshudov, G. N., A. A. Vagin, and E. J. Dodson.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect. D **53:**240–255.
21. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. **276:**307–326.
22. **Posthuma, C. C., D. D. Nedialkova, J. C. Zevenhoven-Dobbe, J. H. Blokhuis, A. E. Gorbalenya, and E. J. Snijder.** 2006. Site-directed mutagenesis of the nidovirus replicative endoribonuclease NendoU exerts pleiotropic effects on the arterivirus life cycle. J. Virol. **80:**1653–1661.
23. **Prentice, E., J. McAuliffe, X. Lu, K. Subbarao, and M. R. Denison.** 2004. Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. J. Virol. **78:**9977–9986.
24. **Qin, L., B. Xiong, C. Luo, Z. M. Guo, P. Hao, J. Su, P. Nan, Y. Feng, Y. X. Shi, X. J. Yu, X. M. Luo, K. X. Chen, X. Shen, J. H. Shen, J. P. Zou, G. P. Zhao, T. L. Shi, W. Z. He, Y. Zhong, H. L. Jiang, and Y. X. Li.** 2003. Identification of probable genomic packaging signal sequence from SARS-CoV genome by bioinformatics analysis. Acta Pharmacol. Sin. **24:**489–496.
25. **Ramjeesingh, M., L. J. Huan, E. Garami, and C. E. Bear.** 1999. Novel method for evaluation of the oligomeric structure of membrane proteins. Biochem. J. **342:**119–123.
26. **Read, R. J.** 2001. Pushing the boundaries of molecular replacement with maximum likelihood. Acta Crystallogr. Sect. D **57:**1373–1382.
27. **Renzi, F., E. Caffarelli, P. Laneve, I. Bozzoni, M. Brunori, and B. Vallone.** 2006. The structure of the endoribonuclease XendoU: from small nucleolar RNA processing to severe acute respiratory syndrome coronavirus replication. Proc. Natl. Acad. Sci. USA **103:**12365–12370.
28. **Ricagno, S., M. P. Egloff, R. Ulferts, B. Coutard, D. Nurizzo, V. Campanacci, C. Cambillau, J. Ziebuhr, and B. Canard.** 2006. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. Proc. Natl. Acad. Sci. USA **103:**11892–11897.
29. **Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science **300:**1394–1399.
30. **Santarsiero, B. D., D. T. Yegian, C. C. Lee, G. Spraggon, J. Gu, D. Scheibe, D. C. Uber, E. W. Cornell, R. A. Nordmeyer, W. F. Kolbe, J. Jin, A. L Jones, J. M. Jaklevic, P. G. Schultz, and R. C. Stevens.** 2002. An approach to rapid protein crystallization using nanodroplets. J. Appl. Crystallogr. **35:**278–281.
31. **Sawicki, S. G., and D. L. Sawicki.** 2005. Coronavirus transcription: a perspective. Curr. Top. Microbiol. Immunol. **287:**31–55.
32. **Severson, W., L. Partin, C. S. Schmaljohn, and C. B. Jonsson.** 1999. Characterization of the Hantaan nucleocapsid protein-ribonucleic acid interaction. J. Biol. Chem. **274:**33732–33739.
33. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. **331:**991–1004.
34. **Snijder, E. J., S. G. Siddell, and A. E. Gorbalenya.** 2005. The order Nidovirales, p. 390–404. *In* B. W. Mahy and V. ter Meulen (ed.), Topley and Wilson's microbiology and microbial infections, vol. 1. Virology. Hodder Arnold, London, England.
35. **Turner, D. R., and P. J. Butler.** 1986. Essential features of the assembly origin of tobacco mosaic virus RNA as studied by directed mutagenesis. Nucleic Acids Res. **14:**9229–9242.
36. **Watanabe, H., K. Nagayama, N. Enomoto, R. Chinzei, T. Yamashiro, N. Izumi, H. Yatsuhashi, T. Nakano, B. H. Robertson, H. Nakasone, H. Sakugawa, and M. Watanabe.** 2003. Chronic hepatitis delta virus infection with genotype IIb variant is correlated with progressive liver disease. J. Gen. Virol. **84:**3275–3289.
37. **Woo, K., M. Joo, K. Narayanan, K. H. Kim, and S. Makino.** 1997. Murine coronavirus packaging signal confers packaging to nonviral RNA. J. Virol. **71:**824–827.
38. **Xu, X., Y. Zhai, F. Sun, Z. Lou, D. Su, Y. Xu, R. Zhang, A. Joachimiak, X. C. Zhang, M. Bartlam, and Z. Rao.** 2006. New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. J. Virol. **80:**7909–7917.
39. **Zhong, W., R. Dasgupta, and R. Rueckert.** 1992. Evidence that the packaging signal for nodaviral RNA2 is a bulged stem-loop. Proc. Natl. Acad. Sci. USA **89:**11146–11150.

Downloaded from http://jvi.asm.org/ on June 6, 2015 by NANYANG TECHNOLOGICAL UNIVERSITY