STRUCTURAL BIOLOGY
ISSN 2059-7983

Structure and oligomerization state of the C-terminal region of the Middle East respiratory syndrome coronavirus nucleoprotein

Thi Hong Van Nguyen,[a,b] Julie Lichière,[a,b] Bruno Canard,[a,b] Nicolas Papageorgiou,[a,b] Sarah Attoumani,[a,b] François Ferron[a,b]*‡ and Bruno Coutard[a,b]*‡

[a]Aix-Marseille Université, AFMB UMR 7257, 13288 Marseilles, France, and [b]CNRS, AFMB UMR 7257, 13288 Marseilles, France. *Correspondence e-mail: francois.ferron@afmb.univ-mrs.fr, bruno.coutard@afmb.univ-mrs.fr

Received 11 July 2018
Accepted 22 October 2018

Edited by R. McKenna, University of Florida, USA

‡ These authors contributed equally.

Keywords: nucleoproteins; Middle East respiratory syndrome coronavirus; MERS-CoV; *Coronaviridae*; X-ray diffraction; SAXS.

PDB reference: C-terminal domain of MERS-CoV nucleocapsid, 6g13

Supporting information: this article has supporting information at journals.iucr.org/d

Middle East respiratory syndrome coronavirus (MERS-CoV) is a human pathogen responsible for a severe respiratory illness that emerged in 2012. Structural information about the proteins that constitute the viral particle is scarce. In order to contribute to a better understanding of the nucleoprotein (N) in charge of RNA genome encapsidation, the structure of the C-terminal domain of N from MERS-CoV obtained using single-crystal X-ray diffraction is reported here at 1.97 Å resolution. The molecule is present as a dimer in the crystal structure and this oligomerization state is confirmed in solution, as measured by additional methods including small-angle X-ray scattering measurements. Comparisons with the structures of the C-terminal domains of N from other coronaviruses reveals a high degree of structural conservation despite low sequence conservation, and differences in electrostatic potential at the surface of the protein.

1. Introduction

Middle East respiratory syndrome coronavirus (MERS-CoV) is the etiological agent responsible for outbreaks of a severe respiratory disease in the Middle East, with sporadic occurrence in Europe, Africa, Asia and America. MERS-CoV belongs to lineage C of the *Betacoronavirus* genus within the *Coronaviridae* family, one of the four viral families (together with *Arteriviridae*, *Mesoniviridae* and *Roniviridae*) that form the *Nidovirales* order. This group of viruses includes the largest-sized RNA genomes known, *i.e.* *Roniviridae* (∼26 kb) and *Coronaviridae* (from 26.3 to 31.7 kb), as well as the medium-sized *Mesoniviridae* (20.2 kb) and the small-sized *Arteriviridae* (12.7–15.7 kb) (Adams *et al.*, 2013). Despite having a positive-sense single-stranded RNA genome, all of these viruses protect their genomes of dissimilar size with a nucleoprotein (N). The 5′ part of the MERS-CoV genome codes for the polyproteins pp1a and pp1ab, which are processed into 16 nonstructural proteins involved in the replication complex. The 3′ part of the genome codes for viral structural proteins and several accessory proteins, which are translated from a set of subgenomic mRNAs. Among these proteins, the nucleoprotein (N) is a multifunctional protein that plays roles in genome encapsidation (reviewed in Chang *et al.*, 2014), RNA synthesis during replication, host translational shutoff and innate immunity disruption (McBride *et*

al., 2014). The coronavirus N is a 40 kDa protein organized into two folded domains that are called the N-terminal domain (NTD) and the C-terminal domain (CTD). They are separated by a disordered region (called LKR) containing a serine/arginine stretch which could regulate the functions of N upon phosphorylation (He *et al.*, 2004; Surjit *et al.*, 2005; Peng *et al.*, 2008; Wu *et al.*, 2009). Despite many studies, the mechanism by which the RNA genome is encapsidated by N has not been fully unraveled. Indeed, the structure of full-length N is not known, probably owing to its flexibility, and the structures of the NTD and the CTD have been explored separately. However, structural and functional data led to a consensus model in which the formation of the ribonucleocapsid (RNP) requires both multi-site interactions of the RNA along N and oligomerization of N (McBride *et al.*, 2014). In this model, dimers of N generated through a CTD–CTD dimerization process represent the basic module of the RNP (Yu *et al.*, 2006; Jayaram *et al.*, 2006). A recent low-resolution electron-microscopy analysis of N from Murine hepatitis virus seems to converge towards a proposal in which N assembles as small oligomers (from dimers to tetramers) and forms a helical structure in complex with the RNA genome (Gui *et al.*, 2017).

The structure of the MERS-CoV NTD has been determined using X-ray crystallography (Papageorgiou *et al.*, 2016). The overall structure is conserved and consists of a five-stranded antiparallel β-sheet core domain with disordered loops of varying length connecting the strands.

Sequence analysis of the C-terminal region of MERS-CoV N suggests that the CTD is 122 amino acids in length (amino acids 239–361 of the MERS-CoV N sequence) followed by an intrinsically disordered region of 50 amino acids. The structures of CTDs from Human coronavirus NL63 (HCoV-NL63), Avian infectious bronchitis virus (IBV), Severe acute respiratory syndrome coronavirus (SARS-CoV) and Murine hepatitis virus (MHV) have been determined. The structures revealed a conserved structural organization in which the CTD forms a dimer with an intermolecular antiparallel four-stranded β-sheet (Yu *et al.*, 2006; Jayaram *et al.*, 2006; Chen *et al.*, 2007; Ma *et al.*, 2010; Szelazek *et al.*, 2017). This β-sheet is surrounded by eight α-helices and 3_{10}-helices. The CTD displays a positively charged area, highlighting that the CTD is a possible RNA-binding module. However, the surface of the charged area can vary from one viral isoform to another, suggesting variability in its RNA-binding affinity (Chen *et al.*, 2007). Several studies support the idea that RNA binding can occur along the N protein, with multiple binding sites acting cooperatively to form a ribonucleocapsid, as reviewed by Chang *et al.* (2014). However, no structure of full-length coronavirus N is available to support this hypothesis, most likely because the recombinant protein is not stable in solution.

In this study, we present the crystallographic structure of the MERS-CoV CTD at a resolution of 1.97 Å. In addition, we report size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALLS) and small-angle X-ray scattering (SAXS) studies of the CTD in order to gain insight into its structural organization in solution.

2. Materials and methods

2.1. Protein production, purification and characterization

The codon-optimized DNA encoding the CTD (amino acids 239–362) of the MERS-CoV nucleoprotein (strain Betacoronavirus England 1, accession No. KC164505) was synthesized by GenScript. The coding sequence of the CTD was amplified by a two-step PCR before cloning into the expression vector pMCOX20A to enable the fusion of a cleavable thioredoxin/6×His tag at the N-terminus (Lantez *et al.*, 2011). The protein was then produced in *Escherichia coli* T7 Express (DE3) cells (New England Biolabs). Expression was induced overnight at 290 K in Terrific Broth medium with 0.5 m*M* IPTG when the $OD_{600\ nm}$ of the culture reached 0.6. The purification of the protein and tag removal was performed under nondenaturing conditions, as described previously (Lantez *et al.*, 2011). The final preparative size-exclusion chromatography (SEC) step was performed in 10 m*M* HEPES, 300 m*M* NaCl pH 7.5. In order to confirm the oligomerization state of the CTD, analytical SEC with online multi-angle laser light scattering, absorbance and refractive-index (MALLS/UV/RI) detectors was carried out on an Alliance 2695 HPLC system (Waters) using a Silica Gel KW802.5 column (Shodex) equilibrated in 10 m*M* HEPES, 150 m*M* NaCl pH 7.5. Measurement was performed using a triple-angle light-scattering detector (miniDAWN TREOS, Wyatt Technology), a quasi-elastic light-scattering instrument (DynaPro, Wyatt Technology) and a differential refractometer (Optilab rEX, Wyatt Technology). Determination of the molecular weight from the MALLS data was performed using the *ASTRA V* software (Wyatt Technology).

2.2. Protein crystallization

Crystallization trials were performed by the sitting-drop vapor-diffusion method at 293 K in a 96-well Swissci plate using a nanodrop-dispensing robot (Mosquito, TTP Labtech) by screening three different ratios of protein (300, 200 and 100 nl) versus a constant volume of precipitant (100 nl). Wide intensive screening conditions were explored and crystals were found in condition No. 10 (0.2 *M* ammonium nitrate, 2.2 *M* ammonium sulfate) from The $AmSO_4$ Suite (Qiagen). Optimal crystallization conditions were obtained by mixing 200 nl protein solution at 11.5 mg ml^{-1} with 100 nl reservoir solution. Crystals appeared within 48 h and were harvested after a week. The crystals of the MERS-CoV CTD were cryoprotected using 100% trimethylamine *N*-oxide (TMAO) prior to flash-cooling in liquid nitrogen at 110 K.

2.3. Data collection and processing

Diffraction data were collected on the ID23 beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble, France using the large in-vacuum PILATUS 6M-F detector (Dectris, Switzerland) at a wavelength of 0.9787 Å with 0.1° oscillation, 0.037 s exposure time and a temperature of 100 K. A total of 1400 images were measured and were subsequently processed using the *autoPROC* toolbox (Vonrhein *et al.*,

2011). The data were integrated using *XDS* (Kabsch, 2010), and were then analysed with *POINTLESS* (Evans, 2006) and scaled with *AIMLESS* (Evans & Murshudov, 2013). The MERS-CoV CTD data were truncated at 1.97 Å resolution to satisfy completeness and $I/\sigma(I) > 2$ statistics. Phasing was performed by molecular replacement using *Phaser* (McCoy *et al.*, 2007). A search model was generated using *CHAINSAW* (Stein, 2008) after first using *Expresso* (Di Tommaso *et al.*, 2011) to identify the structure with the closest sequence homology (46% identity) as that of the SARS CoV CTD, which was obtained by NMR (PDB entry 2jw8; Takeda *et al.*, 2008) and by X-ray diffraction (PDB entry 2cjr; Chen *et al.*, 2007). Refinement was performed by successive and alternate rounds of refinement with *BUSTER* (Bricogne *et al.*, 2017) and model improvement using *Coot* (Emsley *et al.*, 2010). The final model was evaluated using *MolProbity* (Williams *et al.*, 2018).

Surface electrostatics were calculated using *APBS* (Baker *et al.*, 2001). Figures showing structures and alignments were prepared with *UCSF Chimera* (Pettersen *et al.*, 2004) and with *SeaView* (Gouy *et al.*, 2010) and *ESPript*/*ENDscript* (Robert & Gouet, 2014), respectively. The X-ray structure was validated and deposited in the Protein Data Bank with accession code 6g13. Data-collection and refinement statistics are shown in Table 1.

2.4. Small-angle X-ray scattering (SAXS)

SAXS measurements were recorded on beamline BM29 at the ESRF using a PILATUS 1M detector at a sample-to-detector distance of 2.43 m and a wavelength of 0.09919 nm (12.51 keV). SAXS intensity was measured in the range 0.025–5 nm^{-1}. The scattering vector s is defined here as $s = (4\pi/\lambda)\sin\theta$, where λ is the wavelength of the incident radiation in nanometres and θ is half of the angle between the incident and scattered radiation. Measurements were obtained at seven protein concentrations in the range 0.14–8.51 mg ml^{-1}. The protein buffer consisted of 10 m*M* HEPES, 300 m*M* NaCl pH 7.5. All samples were centrifuged for 15 min at 15 000*g* prior to the experiment to minimize the contribution from aggregated particles. Measurements were made for each protein concentration at 4°C by injecting 45 µl protein sample into a 1.8 mm capillary with flow to minimize radiation damage.

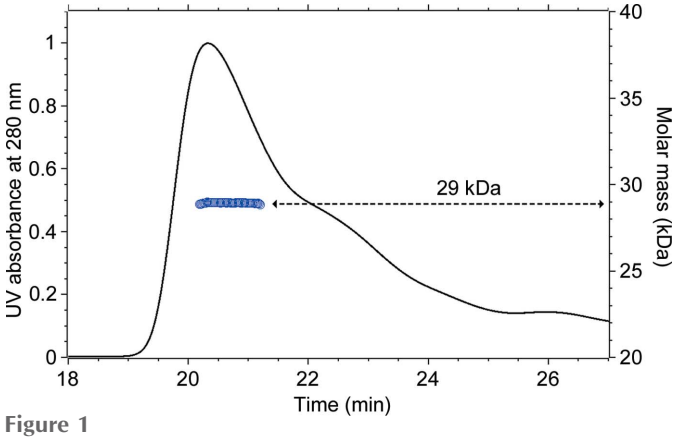

Figure 1
MALLS/UV/refractometry/SEC analysis. The right *y* axis represents the OD at 280 nm and the left *y* axis represents the molar mass in kDa. The absorption peaks are shown as a continuous black line; the blue dots indicate the molar mass.

Table 1
Data processing, structure solution and refinement.

Values in parentheses are for the highest resolution shell.

Data processing	
Wavelength (Å)	0.9787
Space group	$P3_121$
a, *b*, *c* (Å)	120.7, 120.7, 92.6
α, β, γ (°)	90.0, 90.0, 120.0
Resolution range (Å)	69.33–1.97 (2.04–1.97)
Total No. of reflections	110454 (10949)
No. of unique reflections	55298 (5473)
Completeness (%)	100 (99.96)
Multiplicity	2.0 (2.0)
$\langle I/\sigma(I)\rangle$	15.91 (2.47)
R_{meas}	0.03473 (0.4165)
$CC_{1/2}$	0.998 (0.91)
Wilson *B* factor ($Å^2$)	43.9
Structure solution and refinement	
No. of reflections, working set	55294 (5475)
No. of reflections, test set (5.27%)	2912 (303)
R_{cryst}	0.19
R_{free}	0.22
No. of non-H atoms	
Total	3970
Protein	3632
Ligand	28
Water	310
R.m.s. deviations	
Bonds (Å)	0.013
Angles (°)	1.64
Average *B* factors ($Å^2$)	
Overall	64.97
Protein	64.76
Ligand	74.70
Water	66.57
Ramachandran plot	
Favored (%)	98.17
Allowed (%)	1.83
Outliers (%)	0
Clashscore	4.43

Each measurement consisted of ten exposures of 1 s, which were subsequently averaged into a unique intensity profile for each measurement. Buffer reference SAXS measurements were performed before and after the measurement for each protein sample. The forward scattering intensity was calibrated using BSA as a reference at 5 mg ml^{-1}.

Data were processed with the *ATSAS* package (Petoukhov & Svergun, 2007). The ten frames of 1 s exposure were averaged using *PRIMUS* (Konarev *et al.*, 2003). Frames affected by radiation damage were excluded, while the buffer background was subtracted from the sample-scattered intensity signal.

3. Results and discussion

3.1. MERS-CoV CTD behaves as a dimer in solution

Characterization of the purified CTD protein using analytical SEC with online multi-angle laser light scattering (SEC-MALLS) shows that the protein is mainly organized in particles of 29 kDa (elution peak at 20.5 min) corresponding to the molecular weight of a CTD dimer. A second minor population (elution peak at 22.5 min) that was only observed in the analytical SEC and that also displays an apparent

molecular weight (MW) of 29 kDa probably corresponds to the CTD interacting with the silica matrix. Thus, recombinant MERS-CoV CTD is organized in dimers in solution (Fig. 1); this result is consistent with previous observations and with the proposal that N dimerizes through its CTD, forming the core unit leading to assembly of the N polymer (Yu *et al.*, 2006; Jayaram *et al.*, 2006).

To further characterize the oligomerization state of the protein in solution, we studied the protein using SAXS. Fig. 2(*a*) shows the Guinier region $sR_g = 0.1$–1.29 considering a globular type of protein, showing that no trend towards aggregation is present with increasing concentration. Data are shown as open circles and linear regression of data is shown as red lines. For all concentrations we found the same value of radius of gyration R_g, with a mean value (taken over all seven concentrations) of $R_g = 2.0 \pm 0.3$ nm. Mass estimation using the SAXS profile of a standard BSA molecule returns a molecular weight (MW) of 29.32 kDa, which is in agreement with the previous SEC/MALLS measurements. This value is compatible with the MW of a dimer, knowing that the calculated MW of the monomer based on the protein sequence is 14.22 kDa. Again, this result is consistent with previous observations (Yu *et al.*, 2006; Jayaram *et al.*, 2006) and shows that this dimer is not transient but rather stable as it is not significantly affected by low or high protein concentrations.

In order to interpret our data and propose a molecular model, we used a merged data profile truncated between the

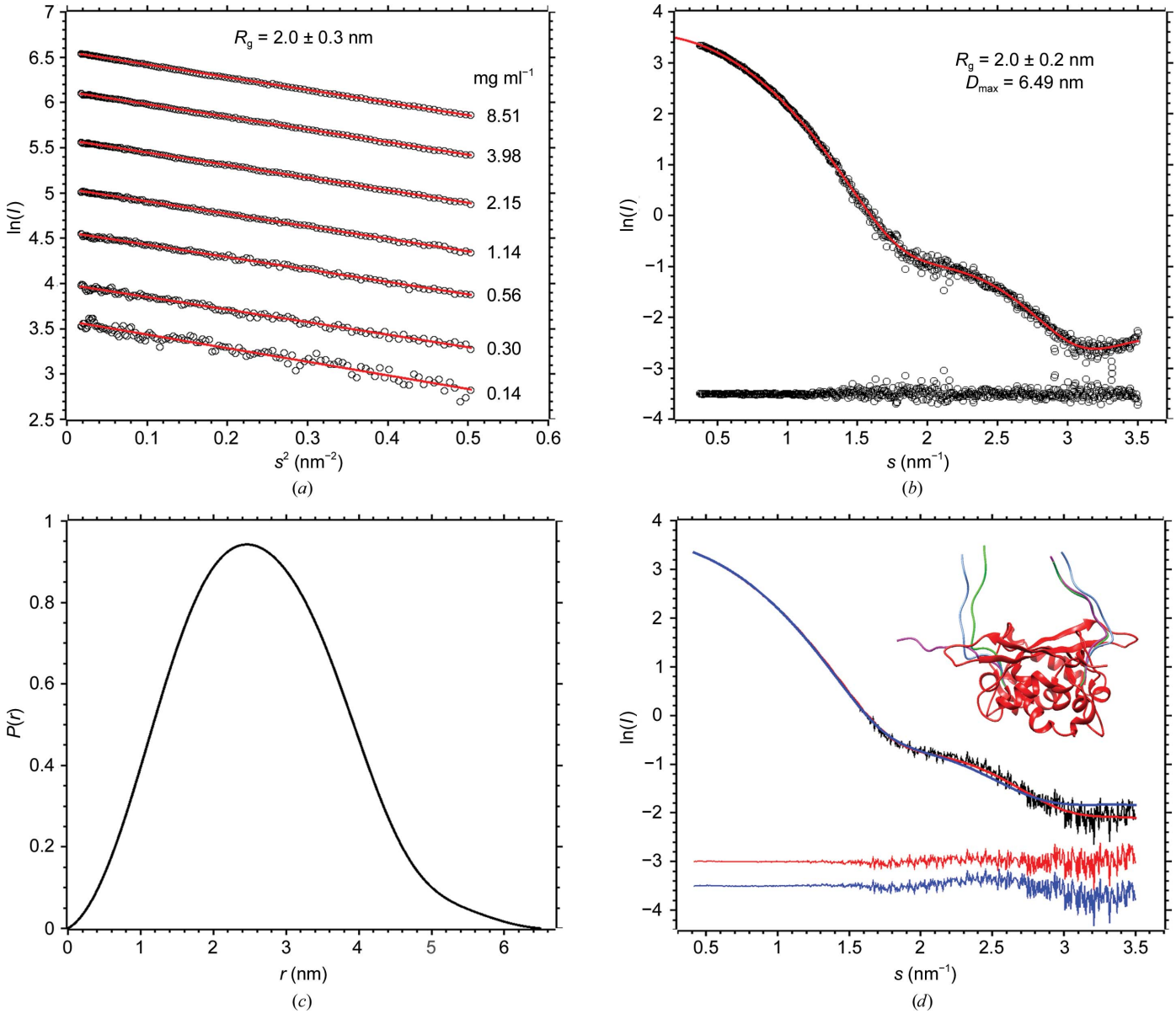

Figure 2
SAXS analysis of the MERS-CoV CTD. (*a*) Guinier plot for all concentrations showing the absence of aggregation. (*b*) *GNOM* fit of the merged profile. (*c*) Distance distribution *P*(*r*) of the merged data. (*d*) *CRYSOL* fit of the model (PDB entry 6g13) with the experimental data (blue line) and *CRYSOL* fit of an enhanced model featuring the disordered N-terminal region residues modeled by *EOM* (red line). The three corresponding enhanced models calculated by *EOM* are shown also in the inset in (*d*).

recommended typical values of the wavevector $s_{min} = \pi/D_{max}$ (0.5 nm^{-1}) and $s_{max} = 7/R_g$ (3.5 nm^{-1}) (Svergun, 1992).

In Fig. 2(*b*) we show the output of the *GNOM* numerical fit over the merged profile, while Fig. 2(*c*) shows the corresponding distance distribution $P(r)$. The latter smoothly converges to zero for a maximum particle dimension of D_{max} = 6.49 nm. The measured dimer dimension of the crystallographic structure that we determined (PDB entry 6g13) is smaller at 5.6 nm. However, ten residues in the N-terminal region of the monomer were not resolved in the crystallographic structure (Thr237–Lys247). These residues contribute to the extra overall length of the molecule. To verify this hypothesis, we modeled these regions using *EOM* analysis (*Ensemble Optimization Method*; Bernardó *et al.*, 2007). Using *Coot*, the ten missing residues were added to the globular part of the molecule and were modeled by *EOM* keeping the globular part fixed. The *EOM* calculations resulted in three possible molecular conformations. *CRYSOL* fits the measured scattering curve with a simulated scattering curve generated using a known X-ray crystal structure. In Fig. 2(*d*) we show two *CRYSOL* fits using the two PDB model structures with and without additional amino acids at the N-terminus. More precisely, the *CRYSOL* fit shown as a blue continuous line corresponds to the crystallographic dimer structure (PDB entry 6g13) with no addition of any supplementary amino acids. The red continuous line is the *CRYSOL* fit corresponding to a mixture of the three molecular conformations calculated by *EOM* after the addition of the missing amino acids to the N-terminal region of each monomer. At the bottom of the figure, the difference between the data and the *CRYSOL* fit is shown. The blue difference spectrum corresponds to the *CRYSOL* fit of PDB entry 6g13, while the red difference spectrum corresponds to the *CRYSOL* fit of the structure calculated by *EOM*. We observe that in the higher resolution range (between 2 and 3.5 nm^{-1}), consideration of the enhanced model calculated by *EOM* improves the fit. This result suggests that the core domain is a dimer in solution, for which a certain flexibility is expected at the N-terminus.

3.2. Structure of the MERS-CoV CTD

A dimer of the MERS-CoV CTD was crystallized. Crystals of an apparent cubic shape with a side length of 150 µm were obtained after 3–5 days. The MERS-CoV CTD crystallized in space group $P3_121$. The asymmetric unit was constituted of two identical dimers rotated by 90° around two axes, with the second dimer being in the center of the first (Fig. 3*a*). The structure was refined to 1.97 Å resolution. The dimer is formed by two molecules related by a noncrystallographic twofold axis and overall looks like a rectangular slab (Fig. 3*b*). An α-helix bundle is on one side of the slab, while the other side is formed by the flat surface of a four-stranded β-sheet flanked by two flexible loops. Each monomeric subunit has an (α+β) fold constituted of five short α-helices and four turns (η) forming a semi-ring, from which two antiparallel β-strands protrude (Fig. 3*c*). It is interesting to note that several residues which are not in contact with adjacent protomers present a high degree of flexibility, presumably leading to poor electron density. Indeed, in the final refined structure we were unable to build several residues of the N-terminus (Thr237–Lys248) or of the loop between the two strands (Asp320–Gly322). This is consistent with the SAXS analysis (see Section 3.1) and shows that the crystallographic structure corresponds to a biological assembly behaving as in solution.

The analysis of the hydrophobic surface of a monomer (Fig. 4) shows the presence of three hydrophobic patches that

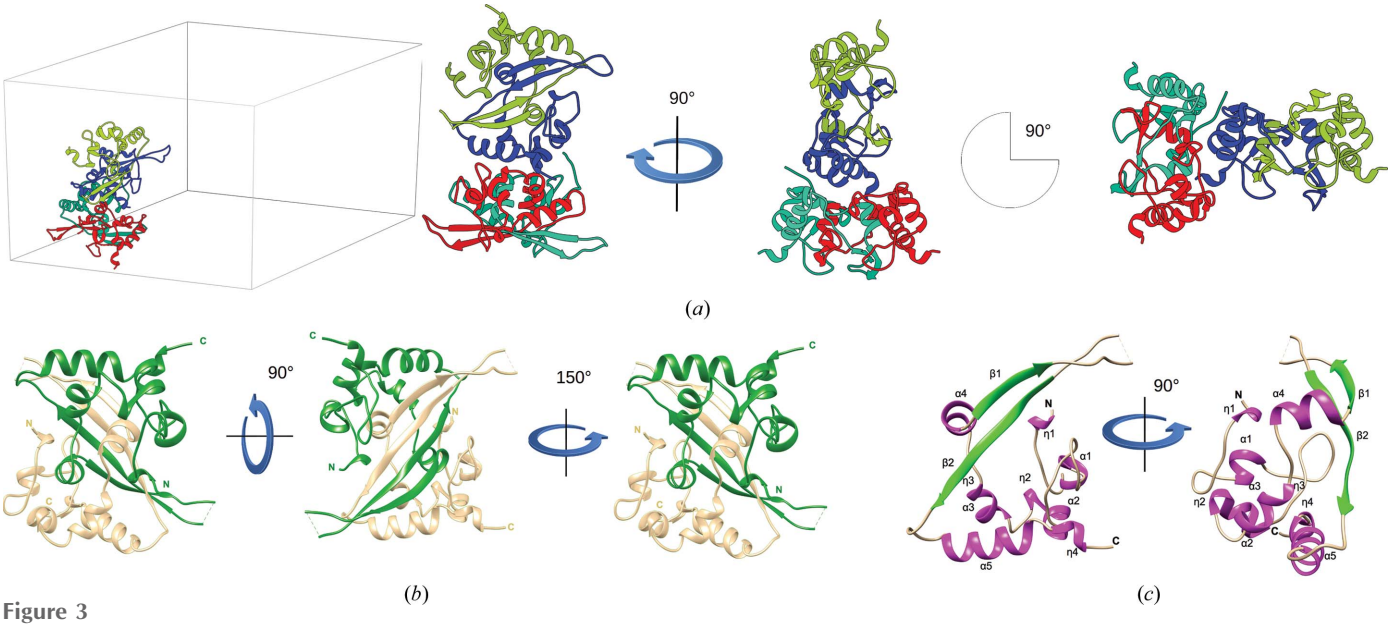

Figure 3
The MERS-CoV CTD structure in different orientations. (*a*) Asymmetric unit illustrating the two dimers at 90° rotation about two axes (dimer 1 is shown in green and red, and dimer 2 is shown in yellow green and blue). (*b*) Clipped monomers shown in ribbon representations and in different orientations, showing how the two chains (green and cream) are intertwined. (*c*) Ribbon representation of the MERS-CoV CTD structure. Helices are colored magenta, strands are colored green and random coils are colored cream; secondary structures are labeled.

interact with the secondary-structure elements of the other protomer. The first patch on the top interacts with helix α5 (Fig. 4), the second patch stacks the core of the two β-strands (Fig. 4) and finally the last patch opposite to the second patch interacts with the N-terminal pseudo-turn η1 (Fig. 4). We can infer from this analysis that the driving force of the dimerization is hydrophobic, giving the overall idea of a 'clipping' mechanism. The core secondary structures of each protomer are tightly intertwined, making this dimer a cornerstone element in the assembly of the ribonucleoprotein, the overall structure of which still needs to be defined. Based on the recent EM analysis (Gui *et al.*, 2017) and prior crystallographic studies on homologous domains (Yu *et al.*, 2006; Jayaram *et al.*, 2006; Ma *et al.*, 2010), this CTD dimer would be the building block of the assembly of a polymeric ribonucleoprotein. Indeed, this assembly is compatible with a head-to-head subunit leading to a helical filament. In this proposed model the growth of the filament is sideways (Ma *et al.*, 2010).

The unique asymmetric unit arrangement observed in the MERS-CoV CTD (Fig. 3*a*), in which the middle of the α-helix bundle of one dimer can interact with the side of the second dimer, suggests that a dimer could prevent the addition of an another 'sideways dimer'. In such a scenario one dimer could act as a 'cap' to prevent further polymerization or regulate the growth of the filament (Supplementary Fig. S1). This type of regulation is reminiscent of the growth control of actin filaments by steric hindrance at one end (Shekhar *et al.*, 2016).

3.3. MERS-CoV CTD sequence and structure conservation

Sequence comparison against the PDB shows that the MERS-CoV CTD has 23–46% identity to homologous domains, including those from IBV, HCoV-NL63 and SARS-CoV. The overall alignment (Supplementary Fig. S2*a*) shows 8.9% identity, highlighting the key residues that are conserved. Considering their positions in the structure, either buried or in loops, we infer that these residues play a structural role to mainly ensure proper folding and/or dimerization (Supplementary Fig. S2*b*). Superposition of the corresponding structures (Supplementary Fig. S3) with the MERS-CoV CTD shows that the folds are virtually identical, with the exception of the flexible ends and the insertion into the β loop.

3.4. MERS-CoV CTD surface electrostatic conservation

In spite of low overall sequence conservation, the properties of the residues are fairly conserved, with an overall similarity of ∼60%. Analysis of the electrostatic surface highlights these similar characteristics between the MERS-CoV, SARS-CoV and IBV CTDs, and almost an inversion of the charges in the HCoV-NL63 CTD. Indeed, in the first three

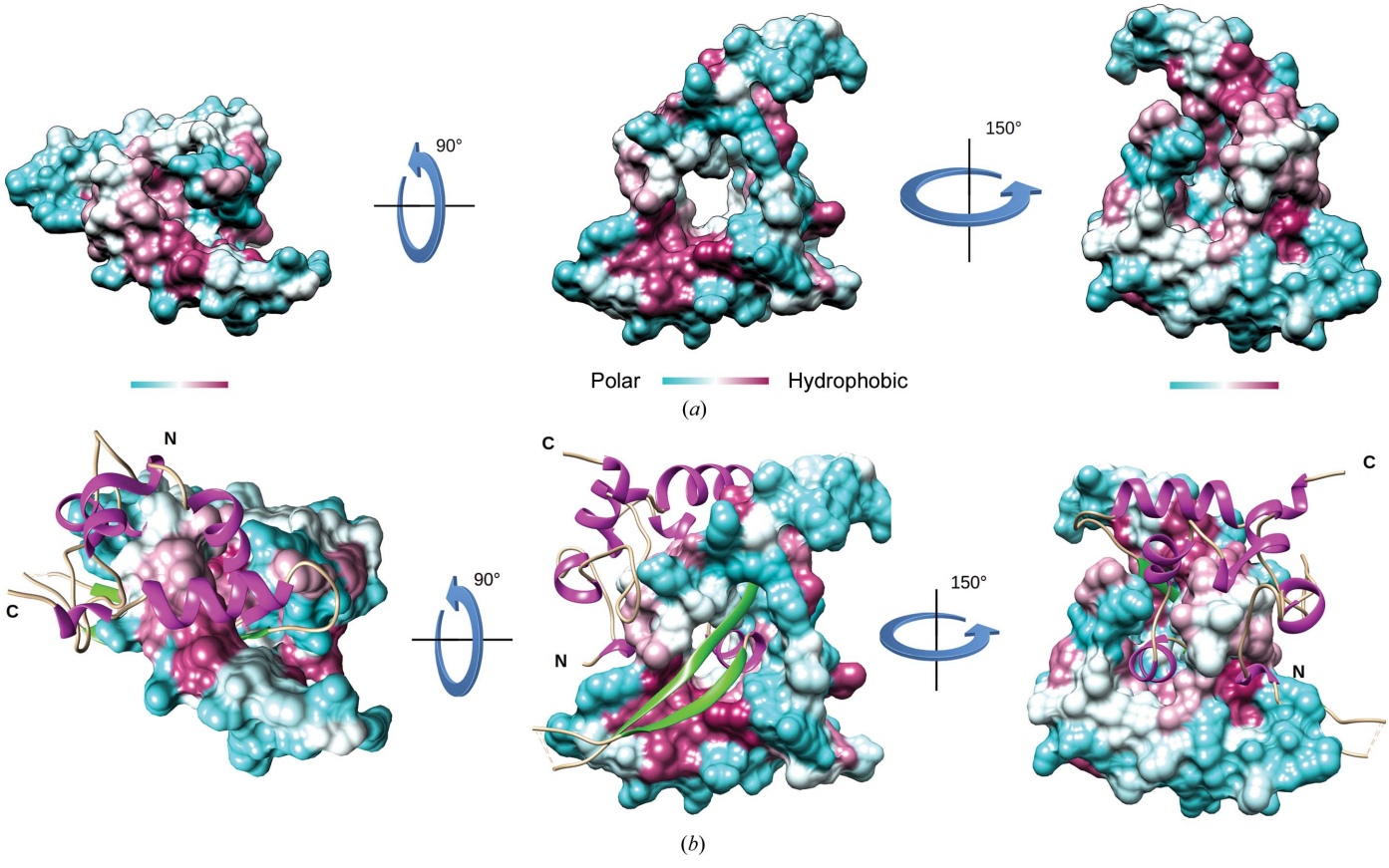

Figure 4
Hydrophobicity analysis of MERS-CoV CTD. (*a*) Surface hydrophobicity showing the hydrophobic potential of a monomer represented by a gradient colored from magenta (high) to cyan (low). (*b*) Surface representation of the MERS-CoV CTD surface hydrophobicity with the second chain of the dimer.

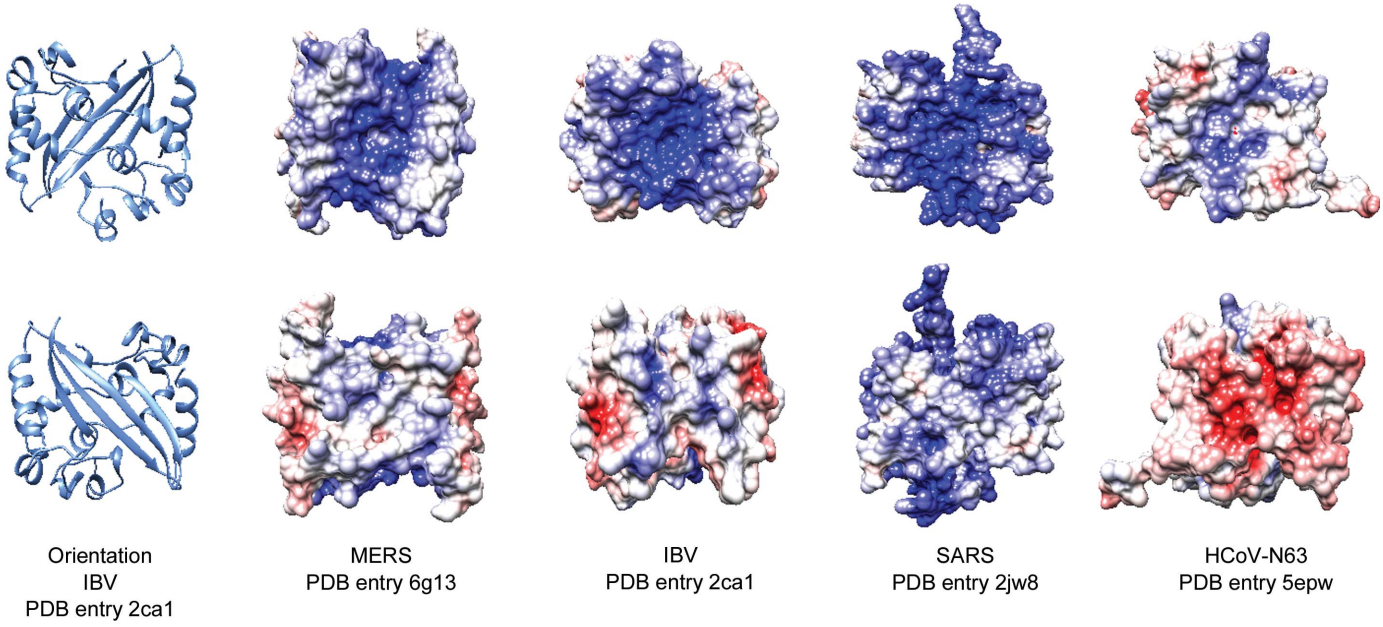

Figure 5
MERS-CoV CTD surface electrostatic comparison with homologous viral domains. The structures presented are from IBV (PDB entry 2ca1; Jayaram *et al.*, 2006), HCoV-NL63 (PDB entry 5epw; Szelazek *et al.*, 2017), SARS-CoV (PDB entry 2cjr; Chen *et al.*, 2007) and MERS-CoV (PDB entry 6g13). The surface electrostatic potential calculated by *APBS* is shown from $-7kT$/e (red) to $+7kT$/e (blue).

dimers the flat side is mostly neutral, while in HCoV-NL63 the charge is negative (Fig. 5, bottom row). On the α-bundle side, a strong positively charged patch is also observed bordering the top and bottom of the dimer, while for the HCoV-63 CTD the bottom and α-bundle sides are either mostly neutral, with a small positive patch in the center of the dimer (Fig. 5, top row). These differences are surprising considering that the sequence of the HCoV-NL63 CTD is not the most distant from that of the MERS-CoV CTD (Supplementary Fig. S2*c*). However, functional studies underscore the role of the CTD in the recognition of the packaging signal in coronaviruses such as Mouse hepatitis virus (Kuo *et al.*, 2014). In particular, MERS-CoV CTD is able to bind a 45-nucleotide stem-loop structure of the packaging signal sequence (Hsin *et al.*, 2018). This observation is consistent with previous observations on SARS-CoV N, where the C-terminal domain alone is able to bind RNA but would rather act in cooperation with other RNA-binding sites along N for formation of the ribonucleocapsid (Chang *et al.*, 2009). In contrast, the CTD of HCoV-NL63, which lacks the positively charged area observed in MERS and SARS-CoV CTD, is not, or is at least poorly, able to bind RNA (Zuwała *et al.*, 2015). Therefore, the structural information described in this study may help to reveal the residues involved in the specific recognition of the packaging signal or possible differences in the mechanisms of RNP assembly.

Aside from its structural role in RNA packaging, MERS-CoV N could also be involved in disease severity by up-regulating the cytokine CXCL 10 (Aboagye *et al.*, 2018). The C-terminal part of N (amino acids 196–413) is sufficient to overexpress CXCL10 and the differences between the charge organization on the CTD in viruses leading to severe syndromes (SARS-CoV and MERS-CoV) and those leading to mild symptoms (HCoV-NL63) could also be investigated in light of the infection outcomes.

Overall, the structural determination of the MERS-CoV CTD reported here allows it to be proposed that the MERS-CoV CTD plays a key role in the formation of the RNP in a similar fashion to other CTDs, and that the CTD might also recruit specific RNA sequences by a mechanism that remains to be elucidated.

Acknowledgements

The authors would like to thank the reviewers for their time and useful comments to improve the manuscript. The authors also thank the ESRF for the provision of synchrotron beam time at BM29 and the ID23-1 staff for support during data collection.

Funding information

This work was supported by the SILVER Large Scale Collaborative Project (grant agreement No. 260644) of the European Union Seventh Framework, the French Infrastructure for Integrated Structural Biology (FRISBI; ANR-10-INSB-05-01) and the Fondation Infection Mediterrannée.

References

Aboagye, J. O., Yew, C. W., Ng, O.-W., Monteil, V. M., Mirazimi, A. & Tan, Y.-J. (2018). *Biosci. Rep.* **38**, BSR20181059.
Adams, M. J., King, A. M. & Carstens, E. B. (2013). *Arch. Virol.* **158**, 2023–2030.
Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. (2007). *J. Am. Chem. Soc.* **129**, 5656–5664.
Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C. & Womack, T. O. (2017). *BUSTER*. Global Phasing Ltd., Cambridge, England.
Chang, C.-K., Hou, M.-H., Chang, C.-F., Hsiao, C.-D. & Huang, T.-H. (2014). *Antiviral Res.* **103**, 39–50.
Chang, C.-K., Hsu, Y.-L., Chang, Y.-H., Chao, F.-A., Wu, M.-C., Huang, Y.-S., Hu, C.-K. & Huang, T.-H. (2009). *J. Virol.* **83**, 2255–2264.
Chen, C.-Y., Chang, C.-K., Chang, Y.-W., Sue, S.-C., Bai, H.-I., Riang, L., Hsiao, C.-D. & Huang, T.-H. (2007). *J. Mol. Biol.* **368**, 1075–1086.
Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J.-M., Taly, J.-F. & Notredame, C. (2011). *Nucleic Acids Res.* **39**, W13–W17.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Gouy, M., Guindon, S. & Gascuel, O. (2010). *Mol. Biol. Evol.* **27**, 221–224.
Gui, M., Liu, X., Guo, D., Zhang, Z., Yin, C.-C., Chen, Y. & Xiang, Y. (2017). *Protein Cell*, **8**, 219–224.
He, R., Leeson, A., Ballantine, M., Andonov, A., Baker, L., Dobie, F., Li, Y., Bastien, N., Feldmann, H., Strocher, U., Theriault, S., Cutts, T., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S. & Li, X. (2004). *Virus Res.* **105**, 121–125.
Hsin, W.-C., Chang, C.-H., Chang, C.-Y., Peng, W.-H., Chien, C.-L., Chang, M.-F. & Chang, S. C. (2018). *J. Biomed. Sci.* **25**, 47.
Jayaram, H., Fan, H., Bowman, B. R., Ooi, A., Jayaram, J., Collisson, E. W., Lescar, J. & Prasad, B. V. (2006). *J. Virol.* **80**, 6612–6620.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.
Kuo, L., Koetzner, C. A., Hurst, K. R. & Masters, P. S. (2014). *J. Virol.* **88**, 4451–4465.
Lantez, V., Dalle, K., Charrel, R., Baronti, C., Canard, B. & Coutard, B. (2011). *PLoS Negl. Trop. Dis.* **5**, e936.
Ma, Y., Tong, X., Xu, X., Li, X., Lou, Z. & Rao, Z. (2010). *Protein Cell*, **1**, 688–697.
McBride, R., van Zyl, M. & Fielding, B. C. (2014). *Viruses*, **6**, 2991–3018.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Papageorgiou, N., Lichière, J., Baklouti, A., Ferron, F., Sévajol, M., Canard, B. & Coutard, B. (2016). *Acta Cryst.* D**72**, 192–202.
Peng, T.-Y., Lee, K.-R. & Tarn, W.-Y. (2008). *FEBS J.* **275**, 4152–4163.
Petoukhov, M. V. & Svergun, D. I. (2007). *Curr. Opin. Struct. Biol.* **17**, 562–571.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
Shekhar, S., Pernier, J. & Carlier, M.-F. (2016). *J. Cell Sci.* **129**, 1085–1091.
Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
Surjit, M., Kumar, R., Mishra, R. N., Reddy, M. K., Chow, V. T. K. & Lal, S. K. (2005). *J. Virol.* **79**, 11476–11486.
Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
Szelazek, B., Kabala, W., Kus, K., Zdzalik, M., Twarda-Clapa, A., Golik, P., Burmistrz, M., Florek, D., Wladyka, B., Pyrc, K. & Dubin, G. (2017). *J. Virol.* **91**, e02503-16.
Takeda, M., Chang, C.-K., Ikeya, T., Güntert, P., Chang, Y.-H., Hsu, Y.-L., Huang, T.-H. & Kainosho, M. (2008). *J. Mol. Biol.* **380**, 608–622.
Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. & Bricogne, G. (2011). *Acta Cryst.* D**67**, 293–302.
Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (2018). *Protein Sci.* **27**, 293–315.
Wu, C.-H., Yeh, S.-H., Tsay, Y.-G., Shieh, Y.-H., Kao, C.-L., Chen, Y.-S., Wang, S.-H., Kuo, T.-J., Chen, D.-S. & Chen, P.-J. (2009). *J. Biol. Chem.* **284**, 5229–5239.
Yu, I.-M., Oldham, M. L., Zhang, J. & Chen, J. (2006). *J. Biol. Chem.* **281**, 17134–17139.
Zuwała, K., Golda, A., Kabala, W., Burmistrz, M., Zdzalik, M., Nowak, P., Kedracka-Krok, S., Zarebski, M., Dobrucki, J., Florek, D., Zeglen, S., Wojarski, J., Potempa, J., Dubin, G. & Pyrc, K. (2015). *PLoS One*, **10**, e0117833.