Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Biochimica et Biophysica Acta 1834 (2013) 1054–1062

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Crystal structure-based exploration of the important role of Arg106 in the RNA-binding domain of human coronavirus OC43 nucleocapsid protein

I.-Jung Chen [a,1], Jeu-Ming P. Yuann [b,1], Yu-Ming Chang [c,1], Shing-Yen Lin [d,1], Jincun Zhao [e], Stanley Perlman [e], Yo-Yu Shen [a], Tai-Huang Huang [f], Ming-Hon Hou [a,d,g,*]

[a] Department of Life Science, National Chung Hsing University, Taichung, 40254, Taiwan
[b] Department of Biotechnology, Ming Chuan University, Taoyuan County, 33348, Taiwan
[c] Institute of Biological Chemistry, Academia Sinica, Taipei, 11529, Taiwan
[d] Institute of Genomics and Bioinformatics, National Chung Hsing University, Taichung, 40254, Taiwan
[e] Department of Microbiology, The University of Iowa, Iowa City, IA 52240, USA
[f] Institute of Biomedical Science, Academia Sinica, Taipei, 115, 11529, Taiwan
[g] Biotechnology Center, National Chung Hsing University, Taichung, 40254, Taiwan

ARTICLE INFO

Article history:
Received 1 October 2012
Received in revised form 6 February 2013
Accepted 5 March 2013
Available online 15 March 2013

Keywords:
Coronaviruses
Nucleocapsid protein
RNA-binding
X-ray crystallography
Arginine106
Virus replication

ABSTRACT

Human coronavirus OC43 (HCoV-OC43) is a causative agent of the common cold. The nucleocapsid (N) protein, which is a major structural protein of CoVs, binds to the viral RNA genome to form the virion core and results in the formation of the ribonucleoprotein (RNP) complex. We have solved the crystal structure of the N-terminal domain of HCoV-OC43 N protein (N-NTD) (residues 58 to 195) to a resolution of 2.0 Å. The HCoV-OC43 N-NTD is a single domain protein composed of a five-stranded β-sheet core and a long extended loop, similar to that observed in the structures of N-NTDs from other coronaviruses. The positively charged loop of the HCoV-OC43 N-NTD contains a structurally well-conserved positively charged residue, R106. To assess the role of R106 in RNA binding, we undertook a series of site-directed mutagenesis experiments and docking simulations to characterize the interaction between R106 and RNA. The results show that R106 plays an important role in the interaction between the N protein and RNA. In addition, we showed that, in cells transfected with plasmids that encoded the mutant (R106A) N protein and infected with virus, the level of the matrix protein gene was decreased by 7-fold compared to cells that were transfected with the wild-type N protein. This finding suggests that R106, by enhancing binding of the N protein to viral RNA plays a critical role in the viral replication. The results also indicate that the strength of N protein/RNA interactions is critical for HCoV-OC43 replication.

1. Introduction

The OC43 strain of the beta coronavirus family (HCoV-OC43), which was first identified in the 1960s, is responsible for ~20% of all "common colds" in humans [1,2]. Although HCoV-OC43 infections are generally mild, more severe upper and lower respiratory tract infections, such as bronchiolitis and pneumonia, have been documented, particularly in infants, elderly individuals, and immunocompromised patients [1,3,4]. Moreover, there have been reports that clusters of HCoV-OC43 infections cause pneumonia in otherwise healthy adults [2,5]. Several studies have also reported that both neurotropism and neuroinvasion of HCoV, particularly the OC43 strain, are associated with multiple sclerosis [6].

CoV particles have an irregular shape defined by an outer envelope with a distinctive club-like shape. Peplomers on the outer envelope give the virus a crown-like (coronal) appearance [7]. The viral genomes of coronaviruses consist of approximately 30 kb of positive sense, single-stranded RNA. These genomes contain several genes encoding structural and nonstructural proteins that are required for the production of progeny virions [1]. The virion envelope that surrounds the nucleocapsid contains the S (spike), M (matrix), and E (envelope) structural proteins. A third glycoprotein, HE (hemagglutinin-esterase), is present in most betacoronaviruses [8,9]. The virion contains a helical nucleocapsid, which consists of the N protein bound to viral RNA. The N protein is the major structural protein of CoVs [10–12]. The formation of the RNP is important for maintaining the RNA in an ordered conformation that is suitable for viral genome replication and transcription [13–16]. Previous studies have shown that the CoV N protein is involved in the regulation of cellular processes, such as gene transcription, actin reorganization, host cell cycle progression, and apoptosis [17–20]. It has also been shown to act as an RNA chaperone [21]. Moreover, the

* Corresponding author at: Biotechnology Center, National Chung Hsing University, Taichung, 40254, Taiwan. Tel.: +886 4 22840338x7011; fax: +886 4 22859329.
E-mail address: mhho@dragon.nchu.edu.tw (M.-H. Hou).
[1] These authors contributed equally to this work.

1570-9639/$ – see front matter.
http://dx.doi.org/10.1016/j.bbapap.2013.03.003

N protein is an important diagnostic marker and is an antigen for the host antibody and T cell immune responses [16,22–24].

Previous studies have revealed that the N- and C-terminal domains of the CoV N proteins, including those of the SARS-CoV, murine hepatitis virus (MHV), and avian infectious bronchitis virus (IBV), are responsible for RNA binding and oligomerization, respectively [25–31]. The central region of the N protein has also been shown to contain an RNA-binding region and primary phosphorylation sites [32–34]. Phosphorylation of the N protein has been shown to play an important role in virus biology [35,36]. To clarify the molecular mechanism of RNP formation in CoVs, the structures of truncated fragments of the N protein, including the N-terminal and C-terminal domains, were investigated [27,37–39]. Despite the conservation of some motifs, the CoV N proteins from various different strains often exhibit different properties, due primarily to their low sequence homology [25].

The N protein of HCoV-OC43, which has a molecular weight of ~50 kDa, is highly basic (pI, 10.0) and exhibits strong hydrophilicity [40]. It also shows only 26–30% amino acid identity to N proteins from other CoV strains [25]. In this study, because of its importance for RNA binding, we chose to characterize the structure of the N-terminus of the HCoV-OC43 N protein using X-ray crystallography. Using the crystal structure and the surface charge distribution of the N-terminus, we were able to further investigate the interactions between the HCoV-OC43 N protein and the RNA molecule. Furthermore, we identified an important role of R106 in the binding of the HCoV-OC43 N protein to RNA using SPR analysis and site-directed mutagenesis. Finally, we present a structural model of the HCoV-OC43 N protein in complex with RNA, which clearly demonstrates the critical role of R106.

2. Materials and methods

All drugs and reagents were purchased from Sigma Chemical Co. (St. Louis, MO) unless otherwise specified. All oligoribonucleotides and oligodeoxyribonucleotides were synthesized using an automated DNA synthesizer and purified by gel electrophoresis. Biotin-linked oligomers were synthesized by incorporating the biotin synthon at the 5′-end; the oligomers were then immobilized to streptavidin-coated biosensor chips for the SPR experiments.

2.1. Expression and purification of the full-length and truncated N proteins

The templates for the HCoV-OC43 N protein were kindly provided by the Institute of Biological Chemistry, Academia Sinica (Taipei, Taiwan). To generate both a full-length versions and the N-terminal domain of the recombinant N protein, the N protein gene was amplified by polymerase chain reaction (PCR) from a plasmid (pGENT) using various primers. The PCR products were digested with *Nde*I and *Xho*I, and the DNA fragments were cloned into pET28a (Novagen) using T4 ligase (NEB). The bacteria transformed with the resultant plasmid were grown in culture. The expression of the recombinant N proteins was induced by supplementing the culture medium with 1 mM IPTG, followed by incubation at 10 °C for 24 h. After harvesting by centrifugation (8,000 g, 10 min, 4 °C), the bacterial pellets were lysed (50 mM Tris-buffered solution at pH 7.3, 150 mM NaCl, 0.1% CHAPS, and 15 mM imidazole). The soluble proteins were obtained from the supernatant following centrifugation (13,000 rpm, 40 min, 4 °C). The methods used for the protein purification have been described previously [41]. Full-length and truncated N proteins carrying a His_6-tag at the N-termini were purified using a Ni-NTA column (Novagen) with an elution gradient ranging from 15 to 300 mM imidazole. Fractions were collected and dialyzed against low-salt buffer. The protein concentrations of the resulting samples were determined using the Bradford method with Bio-Rad protein assay reagents.

2.2. Crystallization

The initial crystallization experiments were set up using Qiagen crystal screens JCSG+ Suite and PACT Suite [42] using the sitting-drop vapor-diffusion method in accordance with our previously described protocol [43]. Each of the crystallization solutions (2 μl) obtained from the screen was mixed with 1.5 μl of purified protein solution (8 mg/ml) and 0.5 μl of 40% hexanediol at room temperature (~298 K) against 400 μl solution in each well of a Cryschem plate. The conditions were refined through seven cycles, and the crystals were grown in a solution containing 0.25 M SPG buffer (pH 6.0) and 25% PEG1500 and then equilibrated at 293 K against 400 μl of the precipitation solution. The SPG buffer was prepared by mixing succinic acid (Sigma), sodium dihydrogen phosphate, (Merck) and glycine (Merck) in a 2:7:7 molar ratio and then adjusting with sodium hydroxide to obtain a pH of 6.0 [44]. The crystals appeared within two weeks, and the largest crystal grew to dimensions of approximately 200 × 100 × 100 μm. The crystals were then soaked in reservoir solution containing 30% (v/v) glycerol as the cryoprotectant prior to being flash-cooled in a nitrogen-gas stream at 100 K. High resolution X-ray data were collected using a synchrotron radiation source. The complete dataset was collected at the beamline BL13B1 in the NSRRC using a ADSC Q315r detector. The crystallographic data integration and reduction were performed using the software package HKL2000 [45]. The crystallographic statistics are listed in Table 1. The Matthews coefficient of 2.06 $Å^3$/Da, which was calculated using *Matthews* (Collaborative Computational Project, 1994) [46], suggested that this structure is likely to represent one molecule in an asymmetric unit. The solvent content was 40.26%. The N-terminal domain of the N protein obtained from the SARS-CoV (PDB ID: 2ofz) was chosen as the initial search model due to its low E-value of 1×10^{-23}. The first molecular replacement trial was performed using the PERON automated interface at the Protein Tectonics Platform (PTP), RIKEN SPring-8 Center, Japan [47]. The best results were obtained using the *MOLREP* program [48]. A single and unambiguous solution for the rotation and translation function was found with the reflections in the resolution range of 3.0–30 Å, a final correlation coefficient of 0.79 and an R factor of 0.44. The structure was refined further using the Crystallography & NMR system (CNS) [49] and deposited in the Protein Data Bank (PDB ID: 4j3k).

2.3. Site-directed mutagenesis

The single mutants (R106A, R106K, R106Q, R106E, R107A, K110A, and R117A) were constructed using a QuikChange™ kit (Stratagene) with a plasmid containing an open reading frame that encodes the full-length HCoV-OC43 N protein as the template for mutagenesis.

Table 1
Crystallographic and refinement data for the HCoV-OC43 N-NTD.

Data collection	
Wavelength (Å)	1.0
Space group	$P6_5$
Unit cell parameter (Å, °)	a = b = 81.57, c = 42.55 α = β = 90°, γ = 120°
Resolution limit (Å)	30–2.0 (2.07–2.00)
Completeness (%)	99.8 (100)
Unique reflections	11,131 (1086)
Redundancy	8.2 (8.5)
R_{merge}	0.037 (0.166)
(I/σ (I))	49.4 (13.77)
Refinement statistics	
R_{cryst} (%)	0.200
R_{free} (%)	0.214
RMSD bond lengths (Å)	0.013
RMSD bond angles (°)	1.978
Most favored region (%)	97.0
Generally allowed region (%)	1.5
Others (%)	1.5
Average B-factor	43.8

The PCR reaction used Pfu DNA polymerase, and each cycle involved heating the sample at 95 °C for 30 s, 55 °C for 1 min, and 68 °C for 2 min/kb of plasmid length; this sequence was repeated for a total of 16 cycles. The templates were digested with *DpnI* and transformed into *E. coli* XL-1 cells. All mutations were confirmed by automated sequencing in both directions.

2.4. Surface plasmon resonance (SPR) binding experiments

The affinity, association, and dissociation of N proteins and RNA were measured in a BIAcore 3000A SPR instrument (Pharmacia, Uppsala, Sweden) equipped with a Sensor Chip SA5 (Pharmacia, Uppsala, Sweden). The apparatus measured binding by monitoring the change in the refractive index of the sensor chip surface. These changes, which were recorded in resonance units (RU), are generally assumed to be proportional to the mass of the molecules bound to the chip. The surface was first washed three times by injecting 10 μl of a solution of 100 mM NaCl and 50 mM NaOH. To control the amount of RNA (or DNA) bound to the SA chip surface, the biotinylated oligomer was immobilized manually onto the surface of a streptavidin chip until a signal of 200RU was achieved in the first cell. The chip surface was then washed with 10 μl of 10 mM HCl to eliminate non-specific binding. The second flow cell was unmodified and served as a control. The appropriate N proteins were dissolved in 50 mM Tris (pH 7.3) with 150 mM NaCl and 0.1% CHAPS and passed over the chip surface for 140 s at a flow rate of 30 μl/min to reach equilibrium. A blank buffer solution was then passed over the chip to initiate the dissociation reaction; this step was continued for an additional 600 s until the reaction was complete. After 600 s, the surface was recovered by washing with 10 μl of 0.1% SDS for each single-stranded RNA. Before fitting to the 1:1 Langmuir model, the binding data were corrected by subtracting the control to account for simple refractive index differences. The sensorgrams of the interactions between RNA and the proteins were analyzed using BIA evaluation software (version 3).

2.5. RNA docking

The computational docking of the ssRNA target onto the HCoV-OC43 N-NTD was performed using HADDOCK1.3 (High Ambiguity Driven DOCKing) software [50]. The starting structures for docking were a model of the single-stranded RNA fragment (5′-UCUAAAC-3′), which was constructed with the DS module Biopolymer, and the crystal structure of the HCoV-OC43 N-NTD. To monitor the interaction between R106 and RNA, R106 was defined as an active residue. The other active residues in the N protein, which included R122, Y124, Y126, and R164, were chosen based on previous studies [51–53] and the solvent accessibility (>50%) determined using the MOL-MOL program. The passive residues of the N protein were defined as the solvent-accessible surface neighbors of the active residues. Default parameters were used for each HADDOCK1.3 run. A total of 1000 docked structures were generated, and the best 100 were refined using simulated annealing. The model structure with the lowest energy was selected for further analysis and is shown in the figures.

2.6. Virus infection and real-time RT-PCR

Real time RT-PCR was performed as previously described [54]. Briefly, 293T cells were cultured in DMEM culture media containing 10% FBS (ATLANTA Biologicals), 1% NEAA (Invitrogen) and 10 μM β-mercaptoethanol (β-ME). A total of 3×10^5 293T cells were seeded into each well of a 12-well plate one day before transfection. The cells were transfected with pcDNA3.1/NP using FuGENE 6 (Roche). After 24 h, the media was removed and then the cells were infected with OC43 virus (MOI = 1) at 33 °C for 2 h in 250 μl DMEM with shaking every 10 min. At day 4 after infection, the media were removed, the cells were lysed in 1 ml Trizol (Invitrogen), RNA was extracted following the manufacturer's instructions, and 2 μg of the RNA was used as the template for cDNA synthesis. cDNA (2 μl) was added to 23 μl of a PCR cocktail that contained 2 × SYBR Green Master Mix (ABI, Foster City, CA) and 0.2 μM of sense and antisense primers (IDT DNA, Coralville, IA). Amplification was performed in an ABI Prism 7700 Thermocycler (ABI). The specificity of the amplification was confirmed using dissociation curve analysis. The data were collected and recorded using the ABI Prism 7700 software and expressed as a function of the threshold cycle (Ct), which is the cycle that the fluorescence intensity in a given reaction tube rises above background (calculated as 10 times the mean standard deviation of the fluorescence in all wells over the baseline cycles). The specific primer sets used for assaying the expression of OC43 matrix (MP) and the housekeeping gene GAPDH were: OC43 MP, Fwd-ATGTTAGGCCGATAA TTGAGGACTAT, Rev-AATGTAAAGATGGCCGCGTAT; GAPDH, Fwd-CCA CTCCTCCACCTTTGA, Rev-ACCCTGTTGCTGTAGCCA.

3. Results and discussion

3.1. The crystal structure of HCoV-OC43 N-NTD

The structure of the HCoV-OC43 N-NTD was determined by the molecular replacement method using the crystal structure of the SARS-CoV N-NTD (PDB ID: 2ofz) as the template. The final protein structure (Fig. 1A) has *R*-factor and *R*-free values of 0.200 and 0.214, respectively

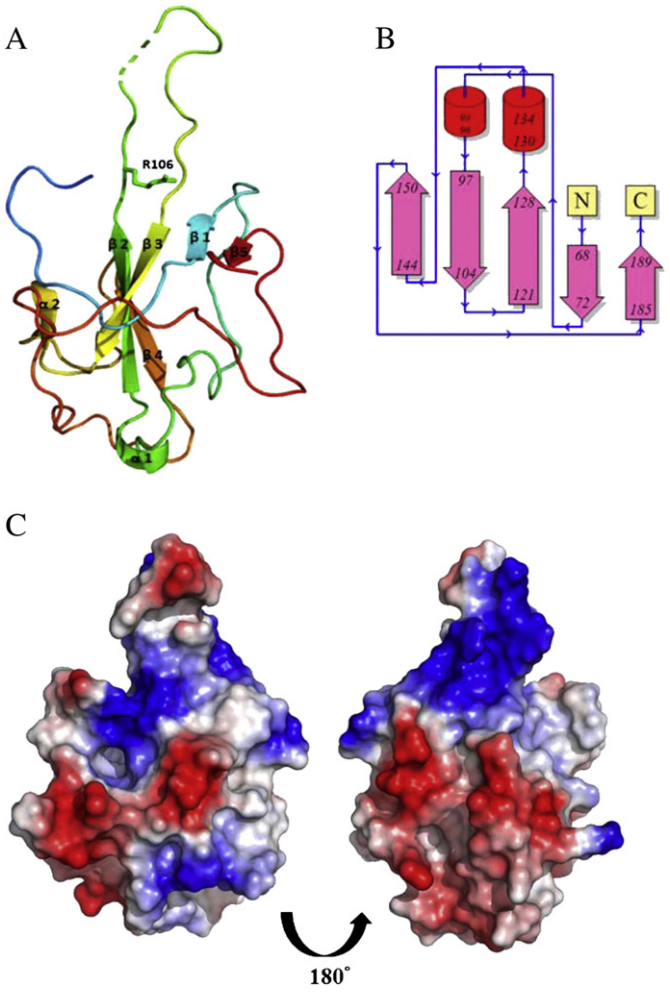

Fig. 1. The structure and topology of the HCoV-OC43 N-NTD. (A) A ribbon diagram of the HCoV-OC43 N-NTD depicts the presence of five β strands, two 3_{10} helices, and several disordered regions. (B) The topology of the HCoV-OC43 N-NTD shows the relative positions of the secondary structures of the truncated protein. (C) The surface charge distribution of the HCoV-OC43 N-NTD.

(Table 1). The core of the structure of HCoV-OC43 N-NTD is a five-stranded anti-parallel β-sheet sandwiched between loops (or short 3_{10} helix) (Fig. 1A and B). Strands β2 and β3 are connected by a long flexible loop composed of amino acid residues 105 to 120 protruding out of the core. However, due to the flexibility of the loop, it was relatively difficult to locate its precise position, particularly within the aa range of 115–117, which is denoted by dotted lines. The MODELLER program was used to model the loop of the HCoV-OC43 N-NTD. The loop region contains four positively charged residues (Arg106, Arg107, Lys110 and Arg117). Based on the surface charge distribution (Fig. 1C), we proposed that the disordered loop of the HCoV-OC43 N-NTD provides a site for binding of the phosphate backbone of RNA through electrostatic interactions.

Previous studies have reported that the N-terminus of the SARS-CoV N protein provides a scaffold for RNA binding [29,37,55,56]. X-ray analysis revealed that the fold of the N-terminal domain of the N protein is

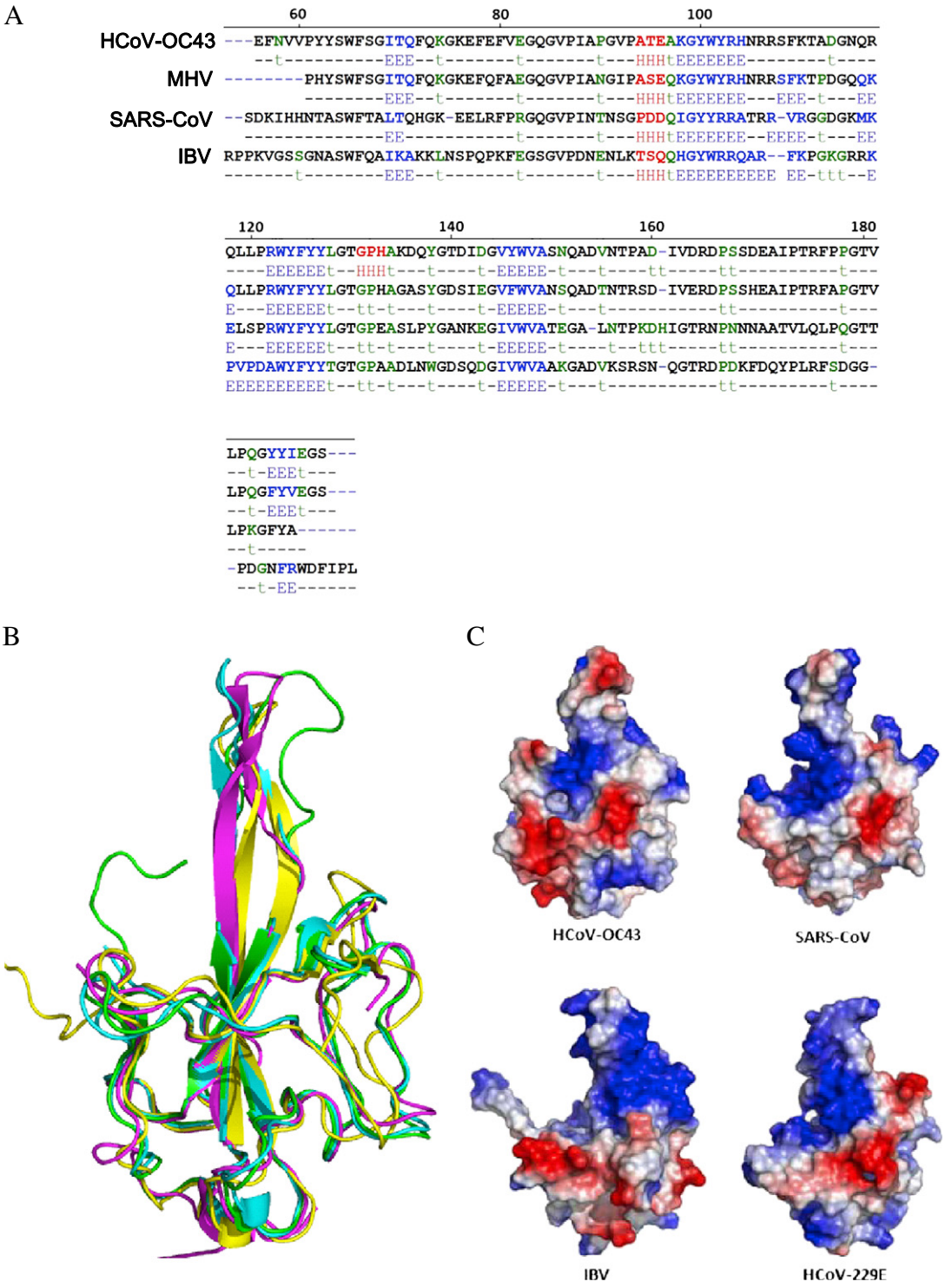

Fig. 2. Amino acid sequences of N-NTDs from various coronaviruses. (A) Secondary structural alignment of the amino acids in N-NTDs from HCoV-OC43, SARS-CoV (PDB ID: 2ofz), IBV (PDB ID: 2gec), and MHV (PDB ID: 3hd4). (B) Superimposition of the HCoV-OC43 N-NTD (green) with N-NTDs from SARS-CoV (magenta), IBV (yellow), and MHV (blue). (C) Surface charge distribution of N-NTDs from HCoV-OC43, SARS-CoV, IBV, and HCoV-229E. The structure of HCoV-229E N-NTD was modeled in the Modeler program using the SARS-CoV N-NTD as a template.

essentially conserved across the various CoVs. It has a right-handed fist-shaped structure, in which palm and finger are rich in basic residues and the flexible loops are organized around the β-sheet core of the N-terminal domain [27,37]. Spencer and Hiscox found that the N-terminal region of the IBV protein facilitates long-range, non-specific interactions between the N protein and viral RNA, thus leading to the formation of the ribonucleocapsid via a lure and lock mechanism [29]. There are multiple RNA binding sites on the N protein and they bind to RNA cooperatively [57]. Similar to the N-terminal domain of the SARS-CoV and IBV N proteins, the N-terminal domain of the HCoV-OC43 N protein was able to bind to RNA. In addition to the N-terminal domain, the highly positively charged portions of the central linker region also show RNA binding affinity, whereas the C-terminal domain of the HCoV-OC43 N protein does not bind RNA [58].

3.2. Analysis and comparison of the structures of HCoV N-NTD

The secondary structure of the HCoV-OC43 N-NTD was compared to those of N-NTDs from other coronaviruses, e.g., MHV, SARS-CoV, and IBV, using the Sequence Annotated by Structure (SAS) website. The differences in the distribution of the secondary structures of these coronaviruses are not obvious, as shown in Fig. 2A [28,37,52],

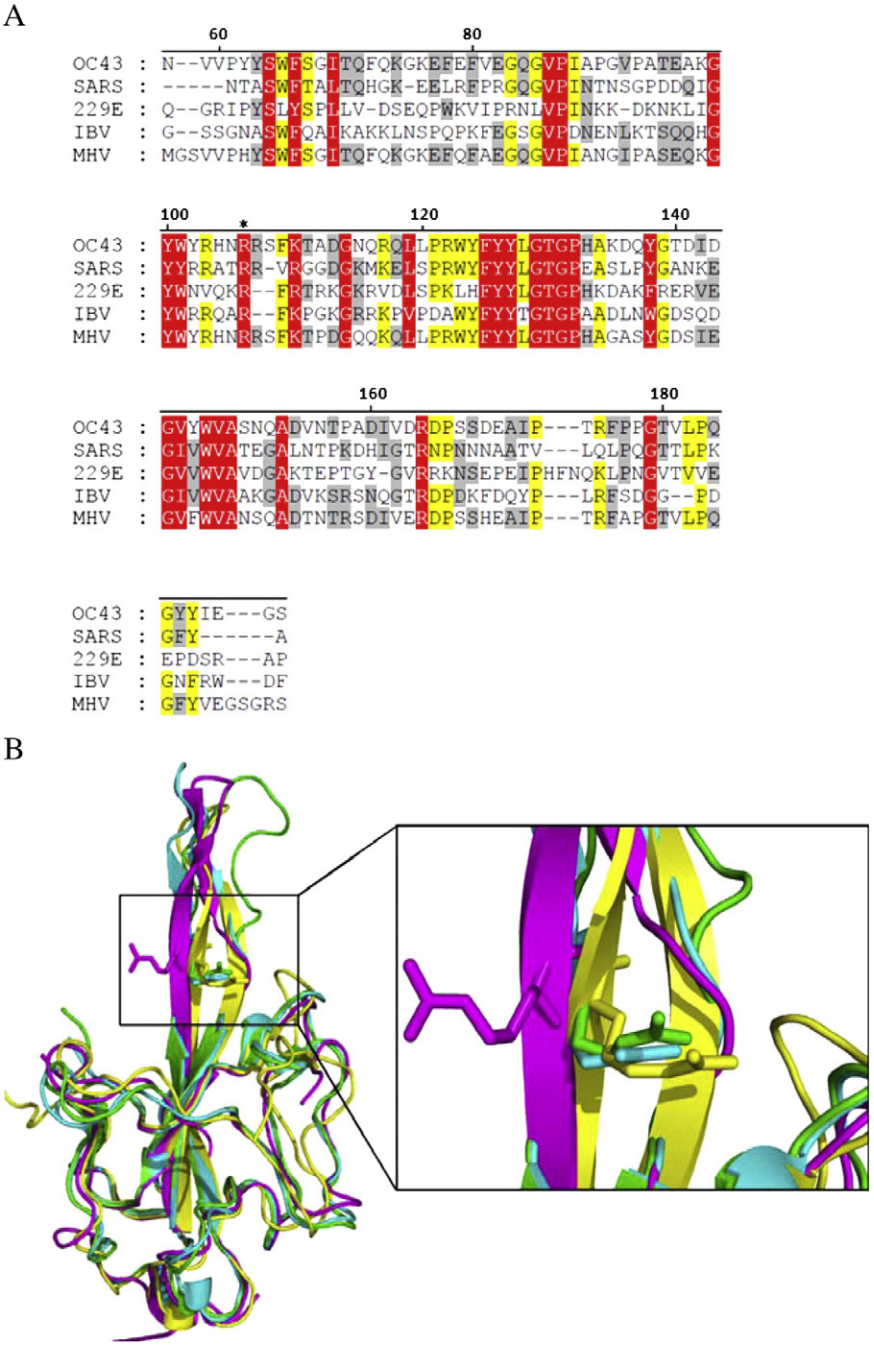

Fig. 3. (A) Amino acid sequence alignment, performed using T-coffee, of N-NTDs from HCoV-OC43 (NP_937954), SARS-CoV (ABI96968), HCoV-229E (AAG48597), IBV (AAB24054), and MHV (ACO72897), all of which were retrieved from GenBank. Residue 106 is indicated with an asterisk. (B) Structural superimposition of the HCoV-OC43 N-NTD (green) with N-NTDs from SARS-CoV (magenta) (PDB ID: 2ofz), IBV (yellow) (PDB ID: 2gec), and MHV (blue) (PDB ID: 3hd4). The inset depicts the location of R106.

which implies that N-NTDs are highly conserved among various coronaviruses. However, superimposition of the structures of the HCoV-OC43, MHV, SARS-CoV, and IBV N-NTDs (Fig. 2B) showed that the structures of the HCoV-OC43 N-NTD are quite similar to those of the MHV, SARS-CoV, and IBV N-NTDs because it contains a protruding loop within the β-sheet core. However, in the MHV, SARS-CoV, and IBV N-NTDs the protruding segment comprised a β-hairpin, whereas a flexible loop predominates in the HCoV-OC43 N-NTD. More significantly, as pointed out by Saikatendu et al., the surface charge distribution, as shown in Fig. 2C, is also quite different for N-NTDs from different species [37]. One common feature, however, is that they all possess positively charged amino acids on or near the protruding loop. NMR studies by Huang et al. have shown that the positively charged loop region is the RNA binding site in SARS-CoV N-NTD [55]. MHV, IBV, and HCoV-OC43 N-NTDs also possess similar positively charged amino acids in this area, thus it is likely that this area may also bind to the negatively charged phosphate backbone of RNA via electrostatic interactions [28,37,52,55]. Nonetheless, the patterns of surface charge distributions of HCoV-OC43, SARS-CoV, IBV, and HCoV-229E differ significantly, suggesting that the RNP packaging may be quite different, although they exhibit short stretches of locally conserved sequences with similar electrostatic sequence profiles (Fig. 2C). In fact, RNP core is self-packaged differently among various morphologically distinct nidoviruses [37].

3.3. RNA binding activity analyses of HCoV-OC43 WT and mutant N proteins

The T-coffee package was used to compare the N-NTD sequences of HCoV-OC43 (NP_937954), SARS-CoV (ABI96968), HCoV-229E (AAG48597), IBV (AAB24054), and MHV (ACO72897). The results revealed that the R106 residue is highly conserved among these coronaviruses (Fig. 3A). In addition, structural alignment demonstrated that R106 is located within a positively charged region and protrudes from the protein surface (Fig. 3B). Therefore, we predicted that R106 would play an indispensible role in the interactions between the N protein and RNA. To explore the role of R106 in the interactions between HCoV-OC43 and RNA, the R106 residue in the full-length N protein was replaced by amino acids with various characteristics, e.g., K, E, M, Q, and A, via site-directed mutagenesis. SPR analyses were then performed to measure the binding affinities between RNA and the HCoV-OC43 N proteins (WT and five mutants). A stretch of single-stranded RNA, 5′-bio(UCUAAAC)$_4$-3′, was immobilized manually on the surface of the SPR chip. This 28-mer RNA, located in the 5′-end non-translated core sequence, was shown to be very important for coronavirus replication and binding to the N protein [59]. Traces from the SPR experiments (Fig. 4A) show the binding capacity of RNA to the WT and mutant N proteins. Comparison of the binding capacities of these proteins (Fig. 4B) showed the binding capacity decreasing in the following order: WT ~ R106K > R106Q ~ R106E > R106M > R106A. These results emphasize the importance of the positively charged R106 in the interaction between the N protein and RNA. The kinetic association and dissociation constants, k_a and k_d, respectively, and the dissociation rate constant, K_d (k_d/k_a), were obtained from analyses of the SPR sensorgrams and the results are listed in Table 2. As shown in Fig. 5A, the k_a values followed the order WT ~ R106K > R106Q ~ R106E > R106M > R106A, with the k_a of R106A smaller than that of the WT by 47.3%. On the other hand, the k_d values increased in the following order: R106A ~ R106M > R106E ~ R106Q > R106K > WT (Fig. 5B) with k_d of R106A being 2.5 times larger than that of WT. Taken together, the K_ds for the HCoV-OC43 WT and 5 mutant N proteins decreased such as R106A > R106M > R106E > R106Q > R106K ~ WT (Table 2). The interaction between R106 of the HCoV-OC43 N-NTD and RNA was modeled. The results of the docking simulation, as shown in Fig. 6, showed that the interaction between R106 and RNA is due to the formation of two hydrogen bonds between the NηH group of R106 and the OH groups on the C2 and C3 of the furanose ring in RNA. The εN of R106 also forms hydrogen bond with 5′-phosphate group of adenine residue in RNA.

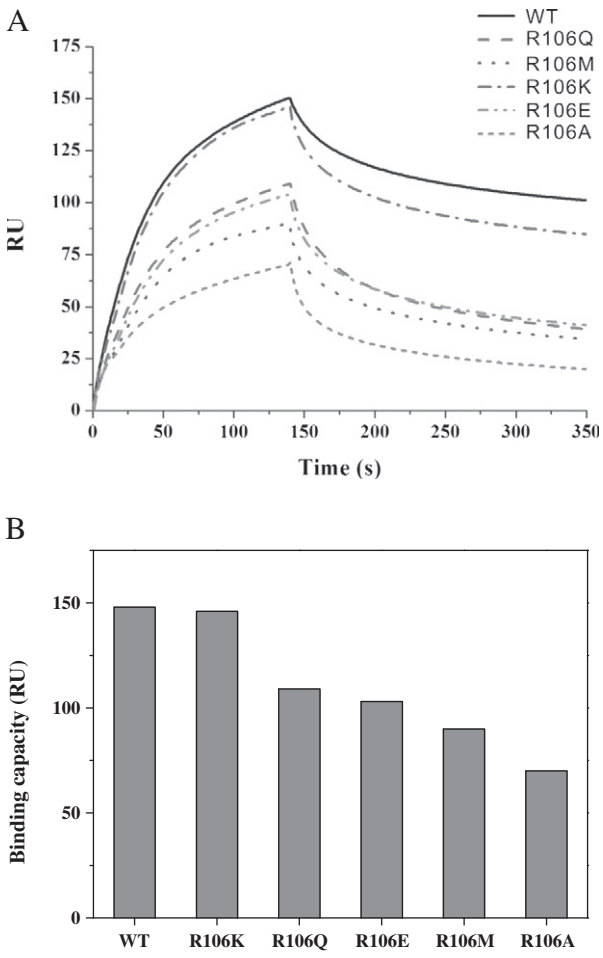

Fig. 4. (A) SPR traces from an RNA binding affinity assay assessing WT HCoV-OC43 N protein and five mutant N proteins. (B) Binding capacity of the WT and mutant N proteins at 0.1 μM for RNA.

To identify the role of R63 in HCoV-229E N proteins, which corresponds to R106 in HCoV-OC43, in RNA-binding, the R63 residue in the HCoV-229E N protein was replaced by valine via site-directed mutagenesis. SPR analyses were then performed to measure the binding affinities between RNA and the HCoV-OC43 N proteins (WT and R63V). The k_a values were 4.03×10^4 and 3.30×10^4 M^{-1} s^{-1} for WT and R63V, respectively. Moreover, the k_d values of WT and R63V were 5.87×10^{-4} and 8.33×10^4 s^{-1}, respectively. The K_d value of the WT was ~2-fold lower than that of R63V. These results suggested

Table 2
Numerical k_a, k_d, and K_d values obtained from the kinetic analysis of the SPR experiments examining binding of HCoV-OC43 WT and mutant N proteins to RNA.

	k_a (M^{-1} s^{-1})/10^4	k_d (s^{-1})/10^{-4}	K_d (nM)
WT	5.01 ± 0.23	5.84 ± 0.31	11.7 ± 0.82
R106A	2.64 ± 0.19	14.8 ± 0.61	56.1 ± 4.64
R106E	3.74 ± 0.28	13.1 ± 0.48	35.0 ± 2.92
R106K	5.37 ± 0.31	7.05 ± 0.32	13.1 ± 0.90
R106M	3.43 ± 0.21	14.6 ± 0.51	42.6 ± 3.01
R106Q	4.06 ± 0.20	11.5 ± 0.45	28.3 ± 1.78
R107A	4.67 ± 0.24	8.81 ± 0.43	18.9 ± 1.33
K110A	5.50 ± 0.22	6.29 ± 0.45	11.4 ± 0.93
R117A	3.24 ± 0.18	5.63 ± 0.34	17.4 ± 1.42

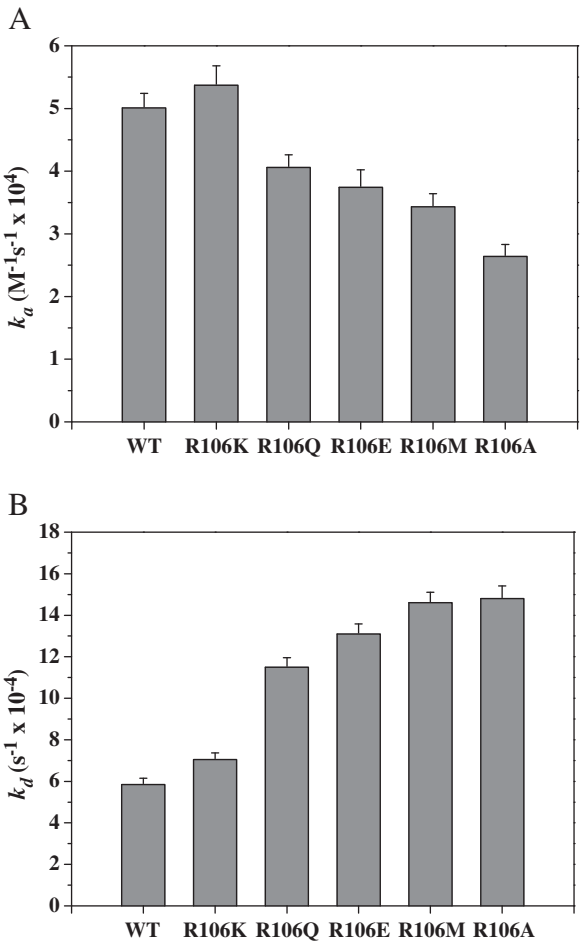

Fig. 5. Kinetic analyses of HCoV-OC43 WT and mutant N proteins binding to RNA: (A) k_a and (B) k_d.

that R63 in HCoV-229E N proteins also play an important role in RNA-binding.

To examine the effects of other conserved positively charged residues in the positively charged loop of N protein including R107, K110

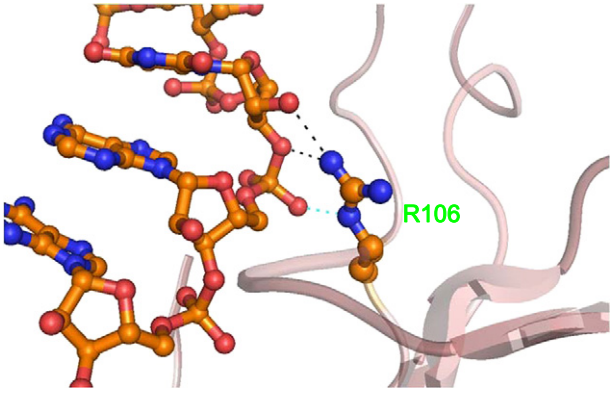

Fig. 6. Predicted interactions between RNA and R106 in the HCoV-OC43 N-NTDs. A single-stranded RNA, 5′-UCUAAAC-3′, was generated using the DS module in biopolymer; the subsequent docking calculations were performed using HADDOCK 1.3 software. The εN of R106 forms hydrogen bond with 5′-phosphate group of adenine residue in RNA (cyan dashed lines). Additionally, the ηN of the Arg106 guanidinium side chain also forms hydrogen bonds with ribose 2′-hydroxyl group and O3′ atom of RNA (black dashed lines). Nitrogen and oxygen atoms were represented by blue ball and red ball, respectively.

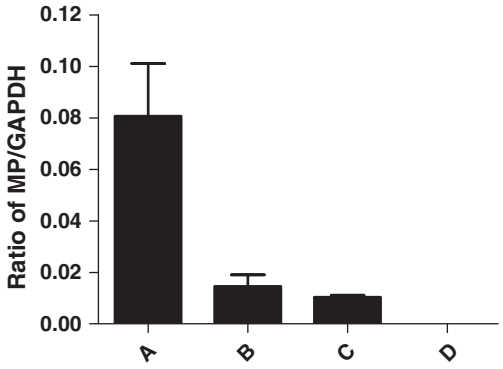

Fig. 7. Virus replication by expression of wild type and R106A N proteins in trans. Cells were transfected with plasmid, pcDNA3.1, encoding (A) the WT N protein or (B) the mutant (R106A) N protein and then infected with HCoV-OC43 as described in Materials and Methods section. Samples were then analyzed for levels of Matrix protein (MP) gene transcript. (C) No transfection and (D) no infection controls are also shown. Quantitative data are shown as mean ± S.D., $n = 3$.

and R117, in RNA-binding, we have created three individual mutants, R107A, K110A, and R117A, of HCoV-OC43 via site-directed mutagenesis and tested their RNA-binding affinity as detailed in this study (Table 2). The k_a values for RNA binding to R107A and R110A are similar to that for WT; the R117A mutant had the smallest k_a value (65%). The k_d values for RNA binding to R117A and R110A are similar to that for WT; the R107A mutant had a larger k_a value (1.5-fold). The K_d values of WT and R110A were essentially the same, while they were increased in R117A and R107A, suggesting that R117 and R107 may participate in RNA-binding of HCoV-OC43 N proteins via electrostatic interactions.

3.4. The effects of R106A mutant on the viral replication

Tan et al. previously demonstrated that R76 in IBV, which corresponds to R106 in HCoV-OC43, plays a functional role on IBV infectivity [53]. In order to determine the effects of R106A mutant on the viral replication, we monitored levels of viral RNA encoding the MP. Since all OC43 genes are transcribed concordantly throughout HCoV-OC43 infection, levels of the MP gene should reflect virus replication. For this purpose, 293T cells were transfected with plasmids encoding the mutant (R106A) N protein and its WT counterpart followed by infection with HCoV-OC43. As shown in Fig. 7, in cells transfected with plasmids encoding the WT N protein and infected with virus, RNA levels of MP were 7-fold increased compared to those detected in cells transfected with plasmids encoding the mutant N protein (R106A), suggesting that expression of WT N protein in trans could stimulate viral replication, and the stimulatory effect of N protein with R106A was obviously impaired. The enhancement of viral replication by expression of CoV N protein in trans was also reported previously [60]. These results support the notion that R106 plays a critical role in virus replication.

4. Conclusion

Coronaviruses (CoVs) cause a wide spectrum of upper and lower respiratory tract infections affecting humans and animals. The strain OC43 identified in 1960s [1,2] and the SARS-CoV identified in 2003 [61] are among the CoVs known to be potential worldwide health risks for human. In CoVs, N protein plays an important role in the packaging of the RNA genome into a helical RNP and in viral RNA synthesis, including replication and transcription. In the present study, we report the crystal structure of the HCoV-OC43 N-NTD, which includes residues 58 to 195. The structure of HCoV-OC43 N-NTD consists of a single-domain five β-sheet structure with a long disordered loops protruding outward, similar to those of the N-NTDs from other coronaviruses. Similar to

that found in IBV [51] by using site-directed mutagenesis, docking stimulation, SPR, and a viral replication assay, we demonstrated that R106 of HCoV-OC43 played a crucial role in RNA binding and virus replication. Therefore, this study could facilitate the development of drugs to disrupt the interaction between the N proteins of CoVs and RNA.

Acknowledgements

We thank Dr. Andrew H.-J. Wang (Academia Sinica) for his help in making this research possible. This work was supported by NSC grant number 100-2113-M-005-004-MY3 (to M.-H. H).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbapap.2013.03.003.

References

[1] J.R. St-Jean, H. Jacomy, M. Desforges, A. Vabret, F. Freymuth, P.J. Talbot, Human respiratory coronavirus OC43: genetic stability and neuroinvasion, J. Virol. 78 (2004) 8824–8834.
[2] A. Vabret, T. Mourez, S. Gouarin, J. Petitjean, F. Freymuth, An outbreak of coronavirus OC43 respiratory infection in Normandy, France, Clin. Infect. Dis. 36 (2003) 985–989.
[3] H.M. El-Sahly, R.L. Atmar, W.P. Glezen, S.B. Greenberg, Spectrum of clinical illness in hospitalized patients with "common cold" virus infections, Clin. Infect. Dis. 31 (2000) 96–100.
[4] A. Gagneur, J. Sizun, S. Vallet, M.C. Legr, B. Picard, P.J. Talbot, Coronavirus-related nosocomial viral respiratory infections in a neonatal and paediatric intensive care unit: a prospective study, J. Hosp. Infect. 51 (2002) 59–64.
[5] R.P. Wenzel, J.O. Hendley, J.A. Davies, J.M. Gwaltney Jr., Coronavirus infections in military recruits. Three-year study with coronavirus strains OC43 and 229E, Am. Rev. Respir. Dis. 109 (1974) 621–624.
[6] N. Arbour, R. Day, J. Newcombe, P.J. Talbot, Neuroinvasion by human respiratory coronaviruses, J. Virol. 74 (2000) 8913–8921.
[7] Q. Leng, Z. Bentwich, A novel coronavirus and SARS, N. Engl. J. Med. 349 (2003) 709.
[8] B.G. Hogue, D.A. Brian, Structural proteins of human respiratory coronavirus OC43, Virus Res. 5 (1986) 131–144.
[9] M.M. Lai, D. Cavanagh, The molecular biology of coronaviruses, Adv. Virus Res. 48 (1997) 1–100.
[10] J.A. Hiscox, RNA viruses: hijacking the dynamic nucleolus, Nat. Rev. Microbiol. 5 (2007) 119–127.
[11] S.A. Stohlman, J.O. Fleming, C.D. Patton, M.M. Lai, Synthesis and subcellular localization of the murine coronavirus nucleocapsid protein, Virology 130 (1983) 527–532.
[12] J.H. You, M.L. Reed, B.K. Dove, J.A. Hiscox, Three-dimensional reconstruction of the nucleolus using meta-confocal microscopy in cells expressing the coronavirus nucleoprotein, Adv. Exp. Med. Biol. 581 (2006) 313–318.
[13] R.S. Baric, G.W. Nelson, J.O. Fleming, R.J. Deans, J.G. Keck, N. Casteel, S.A. Stohlman, Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription, J. Virol. 62 (1988) 4280–4287.
[14] P.S. Masters, M.M. Parker, C.S. Ricard, C. Duchala, M.F. Frana, K.V. Holmes, L.S. Sturman, Structure and function studies of the nucleocapsid protein of mouse hepatitis virus, Adv. Exp. Med. Biol. 276 (1990) 239–246.
[15] K. Pyrc, M.F. Jebbink, B. Berkhout, L. van der Hoek, Genome structure and transcriptional regulation of human coronavirus NL63, Virol. J. 1 (2004) 7.
[16] T.K. Tang, M.P. Wu, S.T. Chen, M.H. Hou, M.H. Hong, F.M. Pan, H.M. Yu, J.H. Chen, C.W. Yao, A.H. Wang, Biochemical and immunological studies of nucleocapsid proteins of severe acute respiratory syndrome and 229E human coronaviruses, Proteomics 5 (2005) 925–937.
[17] L. Du, G. Zhao, Y. Lin, C. Chan, Y. He, S. Jiang, C. Wu, D.Y. Jin, K.Y. Yuen, Y. Zhou, B.J. Zheng, Priming with rAAV encoding RBD of SARS-CoV S protein and boosting with RBD-specific peptides for T cell epitopes elevated humoral and cellular immune responses against SARS-CoV infection, Vaccine 26 (2008) 1644–1651.
[18] P.K. Hsieh, S.C. Chang, C.C. Huang, T.T. Lee, C.W. Hsiao, Y.H. Kou, I.Y. Chen, C.K. Chang, T.H. Huang, M.F. Chang, Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent, J. Virol. 79 (2005) 13848–13855.
[19] S.A. Kopecky-Bromberg, L. Martinez-Sobrido, M. Frieman, R.A. Baric, P. Palese, Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists, J. Virol. 81 (2007) 548–557.
[20] M. Surjit, B. Liu, V.T. Chow, S.K. Lal, The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin–cyclin-dependent kinase complex and blocks S phase progression in mammalian cells, J. Biol. Chem. 281 (2006) 10669–10681.
[21] S. Zuniga, I. Sola, J.L. Moreno, P. Sabella, J. Plana-Duran, L. Enjuanes, Coronavirus nucleocapsid protein is an RNA chaperone, Virology 357 (2007) 215–227.
[22] K.H. Chan, V.C. Cheng, P.C. Woo, S.K. Lau, L.L. Poon, Y. Guan, W.H. Seto, K.Y. Yuen, J.S. Peiris, Serological responses in patients with severe acute respiratory syndrome coronavirus infection and cross-reactivity with human coronaviruses 229E, OC43, and NL63, Clin. Diagn. Lab. Immunol. 12 (2005) 1317–1321.
[23] T. Mourez, A. Vabret, Y. Han, J. Dina, L. Legrand, S. Corbet, F. Freymuth, Baculovirus expression of HCoV-OC43 nucleocapsid protein and development of a Western blot assay for detection of human antibodies against HCoV-OC43, J. Virol. Methods 139 (2007) 175–180.
[24] F.Y. Liang, L.C. Lin, T.H. Ying, C.W. Yao, T.K. Tang, Y.W. Chen, M.H. Hou, Immunoreactivity characterisation of the three structural regions of the human coronavirus OC43 nucleocapsid protein by Western blot: implications for the diagnosis of coronavirus infection, J. Virol. Methods 187 (2013) 413–420.
[25] C.K. Chang, S.C. Sue, T.H. Yu, C.M. Hsieh, C.K. Tsai, Y.C. Chiang, S.J. Lee, H.H. Hsiao, W.J. Wu, W.L. Chang, C.H. Lin, T.H. Huang, Modular organization of SARS coronavirus nucleocapsid protein, J. Biomed. Sci. 13 (2006) 59–72.
[26] C.Y. Chen, C.K. Chang, Y.W. Chang, S.C. Sue, H.I. Bai, L. Riang, C.D. Hsiao, T.H. Huang, Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA, J. Mol. Biol. 368 (2007) 1075–1086.
[27] H. Fan, A. Ooi, Y.W. Tan, S. Wang, S. Fang, D.X. Liu, J. Lescar, The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties, Structure 13 (2005) 1859–1868.
[28] H. Jayaram, H. Fan, B.R. Bowman, A. Ooi, J. Jayaram, E.W. Collisson, J. Lescar, B.V. Prasad, X-ray structures of the N- and C-terminal domains of a coronavirus nucleocapsid protein: implications for nucleocapsid formation, J. Virol. 80 (2006) 6612–6620.
[29] K.A. Spencer, J.A. Hiscox, Characterisation of the RNA binding properties of the coronavirus infectious bronchitis virus nucleocapsid protein amino-terminal region, FEBS Lett. 580 (2006) 5993–5998.
[30] Y.S. Lo, S.Y. Lin, S.M. Wang, C.T. Wang, Y.L. Chiu, T.H. Huang, M.H. Hou, Oligomerization of the carboxyl terminal domain of the human coronavirus 229E nucleocapsid protein, FEBS Lett. 587 (2013) 120–127.
[31] S.M. Kuo, H.W. Kao, M.H. Hou, C.H. Wang, S.H. Lin, H.L. Su, Evolution of infectious bronchitis virus in Taiwan: positively selected sites in the nucleocapsid protein and their effects on RNA-binding activity, Vet. Microbiol. 162 (2013) 408–418.
[32] P.S. Masters, Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus, Arch. Virol. 125 (1992) 141–160.
[33] G.W. Nelson, S.A. Stohlman, Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein, J. Gen. Virol. 74 (Pt 9) (1993) 1975–1979.
[34] D. Peng, C.A. Koetzner, T. McMahon, Y. Zhu, P.S. Masters, Construction of murine coronavirus mutants containing interspecies chimeric nucleocapsid proteins, J. Virol. 69 (1995) 5475–5484.
[35] H. Chen, A. Gill, B.K. Dove, S.R. Emmett, C.F. Kemp, M.A. Ritchie, M. Dee, J.A. Hiscox, Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance, J. Virol. 79 (2005) 1164–1179.
[36] K.A. Spencer, M. Dee, P. Britton, J.A. Hiscox, Role of phosphorylation clusters in the biology of the coronavirus infectious bronchitis virus nucleocapsid protein, Virology 370 (2008) 373–381.
[37] K.S. Saikatendu, J.S. Joseph, V. Subramanian, B.W. Neuman, M.J. Buchmeier, R.C. Stevens, P. Kuhn, Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-terminal domain of N protein, J. Virol. 81 (2007) 3913–3921.
[38] M. Takeda, C.K. Chang, T. Ikeya, P. Guntert, Y.H. Chang, Y.L. Hsu, T.H. Huang, M. Kainosho, Solution structure of the c-terminal dimerization domain of SARS coronavirus nucleocapsid protein solved by the SAIL-NMR method, J. Mol. Biol. 380 (2008) 608–622.
[39] I.M. Yu, M.L. Oldham, J. Zhang, J. Chen, 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 (2006) 17134–17139.
[40] T. Kamahora, L.H. Soe, M.M. Lai, Sequence analysis of nucleocapsid gene and leader RNA of human coronavirus OC43, Virus Res. 12 (1989) 1–9.
[41] H.C. Hung, C.L. Liu, J.T. Hsu, J.T. Horng, M.Y. Fang, S.Y. Wu, S.H. Ueng, M.Y. Wang, C.W. Yaw, M.H. Hou, Development of an anti-influenza drug screening assay targeting nucleoproteins with tryptophan fluorescence quenching, Anal. Chem. 84 (2012) 6391–6399.
[42] J. Newman, Expanding screening space through the use of alternative reservoirs in vapor-diffusion experiments, Acta Crystallogr. D Biol. Crystallogr. 61 (2005) 490–493.
[43] I.J. Chen, C.C. Chou, C.L. Liu, C.C. Lee, L.S. Kan, M.H. Hou, Crystallization and preliminary X-ray diffraction analysis of the N-terminal domain of human coronavirus OC43 nucleocapsid protein, Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66 (2010) 815–818.
[44] J. Newman, Novel buffer systems for macromolecular crystallization, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 610–612.
[45] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276 (1997) 307–326.
[46] B.W. Matthews, Solvent content of protein crystals, J. Mol. Biol. 33 (1968) 491–497.
[47] M. Sugahara, Y. Asada, K. Shimizu, H. Yamamoto, N.K. Lokanath, H. Mizutani, B. Bagautdinov, Y. Matsuura, M. Taketa, Y. Kageyama, N. Ono, Y. Morikawa, Y. Tanaka, H. Shimada, T. Nakamoto, M. Sugahara, M. Yamamoto, N. Kunishima, High-throughput crystallization-to-structure pipeline at RIKEN SPring-8 Center, J. Struct. Funct. Genomics 9 (2008) 21–28.
[48] A. Vagin, A. Teplyakov, Molecular replacement with MOLREP, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 22–25.

[49] A.T. Brunger, 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, G.L. Warren, Crystallography & NMR system: a new software suite for macromolecular structure determination, Acta Crystallogr. D Biol. Crystallogr. 54 (1998) 905–921.
[50] C. Dominguez, R. Boelens, A.M. Bonvin, HADDOCK: a protein–protein docking approach based on biochemical or biophysical information, J. Am. Chem. Soc. 125 (2003) 1731–1737.
[51] S.C. Keane, P. Liu, J.L. Leibowitz, D.P. Giedroc, Functional transcriptional regulatory sequence (TRS) RNA binding and helix destabilizing determinants of murine hepatitis virus (MHV) nucleocapsid (N) protein, J. Biol. Chem. 287 (2012) 7063–7073.
[52] N.E. Grossoehme, L. Li, S.C. Keane, P. Liu, C.E. Dann III, J.L. Leibowitz, D.P. Giedroc, Coronavirus N protein N-terminal domain (NTD) specifically binds the transcriptional regulatory sequence (TRS) and melts TRS-cTRS RNA duplexes, J. Mol. Biol. 394 (2009) 544–557.
[53] Y.W. Tan, S. Fang, H. Fan, J. Lescar, D.X. Liu, 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 (2006) 4816–4825.
[54] L. Pewe, H. Zhou, J. Netland, C. Tangudu, H. Olivares, L. Shi, D. Look, T. Gallagher, S. Perlman, A severe acute respiratory syndrome-associated coronavirus-specific protein enhances virulence of an attenuated murine coronavirus, J. Virol. 79 (2005) 11335–11342.
[55] Q. Huang, L. Yu, A.M. Petros, A. Gunasekera, Z. Liu, N. Xu, P. Hajduk, J. Mack, S.W. Fesik, E.T. Olejniczak, Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein, Biochemistry 43 (2004) 6059–6063.
[56] Y. Ma, X. Tong, X. Xu, X. Li, Z. Lou, Z. Rao, Structures of the N- and C-terminal domains of MHV-A59 nucleocapsid protein corroborate a conserved RNA-protein binding mechanism in coronavirus, Protein Cell 1 (2010) 688–697.
[57] C.K. Chang, Y.L. Hsu, Y.H. Chang, F.A. Chao, M.C. Wu, Y.S. Huang, C.K. Hu, T.H. Huang, Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: implications for ribonucleocapsid protein packaging, J. Virol. 83 (2009) 2255–2264.
[58] C.Y. Huang, Y.L. Hsu, W.L. Chiang, M.H. Hou, Elucidation of the stability and functional regions of the human coronavirus OC43 nucleocapsid protein, Protein Sci. 18 (2009) 2209–2218.
[59] G.W. Nelson, S.A. Stohlman, S.M. Tahara, High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA, J. Gen. Virol. 81 (2000) 181–188.
[60] J. Pan, X. Peng, Y. Gao, Z. Li, X. Lu, Y. Chen, M. Ishaq, D. Liu, M.L. Dediego, L. Enjuanes, D. Guo, Genome-wide analysis of protein-protein interactions and involvement of viral proteins in SARS-CoV replication, PLoS One 3 (2008) e3299.
[61] J.S. Peiris, C.M. Chu, V.C. Cheng, K.S. Chan, I.F. Hung, L.L. Poon, K.I. Law, B.S. Tang, T.Y. Hon, C.S. Chan, K.H. Chan, J.S. Ng, B.J. Zheng, W.L. Ng, R.W. Lai, Y. Guan, K.Y. Yuen, Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study, Lancet 361 (2003) 1767–1772.