Journal of Virology, Mar. 2008, p. 2515–2527 Vol. 82, No. 5
0022-538X/08/$08.00+0 doi:10.1128/JVI.02114-07
Copyright © 2008, American Society for Microbiology. All Rights Reserved.

Structures of Two Coronavirus Main Proteases: Implications for Substrate Binding and Antiviral Drug Design[▽]

Xiaoyu Xue,[1,2]† Hongwei Yu,[3]† Haitao Yang,[1,2] Fei Xue,[1,2] Zhixin Wu,[3] Wei Shen,[1,2] Jun Li,[1,2]
Zhe Zhou,[1] Yi Ding,[1] Qi Zhao,[1,2] Xuejun C. Zhang,[2] Ming Liao,[3]
Mark Bartlam,[1,2,4] and Zihe Rao[1,2,4]*

Tsinghua-Nankai-IBP Joint Research Group for Structural Biology, Tsinghua University, Beijing 100084, China[1]*;*
National Laboratory of Biomacromolecules, Institute of Biophysics (IBP), Chinese Academy of Sciences,
Beijing 100101, China[2]*; Laboratory of Avian Medicine, College of Veterinary Medicine,*
South China Agricultural University, Guangzhou 510642, China[3]*; and*
College of Life Sciences, Nankai University, Tianjin 300071, China[4]

Received 25 September 2007/Accepted 12 December 2007

Coronaviruses (CoVs) can infect humans and multiple species of animals, causing a wide spectrum of diseases. The coronavirus main protease (M^{pro}), which plays a pivotal role in viral gene expression and replication through the proteolytic processing of replicase polyproteins, is an attractive target for anti-CoV drug design. In this study, the crystal structures of infectious bronchitis virus (IBV) M^{pro} and a severe acute respiratory syndrome CoV (SARS-CoV) M^{pro} mutant (H41A), in complex with an N-terminal autocleavage substrate, were individually determined to elucidate the structural flexibility and substrate binding of M^{pro}. A monomeric form of IBV M^{pro} was identified for the first time in CoV M^{pro} structures. A comparison of these two structures to other available M^{pro} structures provides new insights for the design of substrate-based inhibitors targeting CoV M^{pro}s. Furthermore, a Michael acceptor inhibitor (named N3) was cocrystallized with IBV M^{pro} and was found to demonstrate in vitro inactivation of IBV M^{pro} and potent antiviral activity against IBV in chicken embryos. This provides a feasible animal model for designing wide-spectrum inhibitors against CoV-associated diseases. The structure-based optimization of N3 has yielded two more efficacious lead compounds, N27 and H16, with potent inhibition against SARS-CoV M^{pro}.

Coronaviruses (CoVs) are highly prevalent and severe pathogens that cause a wide range of diseases in multiple species of animals, including humans (16, 25, 30, 36). In 2003, the etiological agent responsible for the global outbreak of a life-threatening atypical pneumonia that caused approximately 800 deaths worldwide was identified as the severe acute respiratory syndrome CoV (SARS-CoV) (7, 9, 14, 15, 24). A prototype of the *Coronaviridae* family is avian infectious bronchitis virus (IBV) (16, 30), which belongs to the genetic group III of CoV (16) and causes considerable economic losses for the poultry industry worldwide (5, 13).

CoVs are enveloped positive-stranded RNA viruses with the largest viral RNA genomes known to date, ranging from 27 to 31 kb (16). The CoV replicase gene encodes two overlapping polyproteins, termed pp1a and pp1ab, which mediate viral replication and transcription (3, 16, 29, 36). The maturation of CoVs involves a highly complex cascade of proteolytic processing events on the polyproteins to control viral gene expression and replication. Most maturation cleavage events within the precursor polyprotein are mediated by the CoV main protease (CoV M^{pro}; also known as 3CL protease or $3CL^{pro}$), a three-domain (domains I to III) cysteine protease with a chymotrypsin-like two-domain fold at the N terminus (10, 18, 37). The structures of CoV M^{pro}s revealed that two CoV M^{pro} molecules form an active homodimer (1, 2, 33, 35). A Cys-His catalytic dyad is located in a cleft between domains I and II (1, 2, 35), and the N-terminal residues 1 to 7 (or N finger) of M^{pro} are considered to play an important role in the proteolytic activity (1, 2, 33, 35). The C-terminal domain III is reported to be required for dimerization (28).

Here, we report the crystal structures of two CoV M^{pro}s. The first is the IBV M^{pro} structure with a dimeric form and a unique monomeric form in one asymmetric unit. The monomeric form has not been observed in any of the previously reported CoV M^{pro}s; its C terminus inserts into one of the active sites present in the dimer. The second is the structure of an active-site mutant, H41A, of SARS-CoV M^{pro} in complex with the N-terminal 11-amino-acid peptide as the substrate, which provides insights into the substrate binding and specificity of the S1′ to S5′ sites in SARS-CoV M^{pro} in an unprecedented way.

As the CoV M^{pro} is responsible for the maturation of itself and the subsequent maturation of the replicase polyproteins (37), it has become an attractive target for anti-CoV drug design. Here, we also present the cocrystal structure of IBV M^{pro} in complex with N3, a wide-spectrum inhibitor that we designed previously to target CoV M^{pro}s (34). We further demonstrate its rapid in vitro inactivation against the viral protease and potent antiviral activity toward IBV in chicken embryos. This assay provides an easily accessible animal model for optimizing wide-spectrum inhibitors against CoV-associated diseases. A comparison of the substrate binding sites of

---

* Corresponding author. Mailing address: Laboratory of Structural Biology, Life Sciences Building, Tsinghua University, Beijing 100084, China. Phone: 86 10 62771493. Fax: 86 10 62773145. E-mail: raozh@xtal.tsinghua.edu.cn.

† These authors contributed equally.

▽ Published ahead of print on 19 December 2007.

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

IBV M^{pro} and SARS-CoV M^{pro} provides further insights for the design of substrate-based inhibitors targeting CoV M^{pro}s. Further modification of Michael acceptor inhibitors based on the new structural information provided here results in two improved inhibitors, termed N27 and H16, with potent inhibition against SARS-CoV M^{pro}.

MATERIALS AND METHODS

Protein purification and crystallization. The protein expression, purification, and crystallization of native IBV M^{pro} has been described previously (20, 34). The crystal structure of IBV M^{pro} could not be determined using conventional molecular replacement techniques. Therefore, a selenomethionyl (SeMet) derivative of IBV M^{pro} was prepared for crystallization and data collection. The recombinant plasmid pGEX-4T-1-IBV M^{pro} was used to transform the methionine auxotrophic B834 (DE3) *Escherichia coli* strain (Novagen), which was propagated in minimal medium supplemented with 60-mg liter^{-1} L-SeMet. The SeMet-substituted IBV M^{pro} was purified as described before and concentrated to 20 mg ml^{-1} for crystallization. The best crystals were obtained using streak seeding, with 2.5% (wt/vol) polyethylene glycol 4000 (PEG4K), 12% (vol/vol) 2-propanol, and 0.1 M sodium cacodylate (pH 6.5) as the reservoir solution.

Crystals of IBV M^{pro} complexed with inhibitor N3 were produced by cocrystallization. IBV M^{pro} was incubated with an equal molar concentration of N3 for 24 h at 4°C. This complex did not crystallize under conditions described above. However, single cubic crystals were obtained in 1 day by the hanging drop vapor diffusion method at 18°C using a reservoir solution containing 20% (wt/vol) PEG10K and 0.1 M HEPES (pH 7.5) without any seeds.

The coding sequence of SARS-CoV M^{pro} was cloned from the SARS-CoV BJ01 strain and inserted into the BamHI and XhoI sites of pGEX-6p-1 plasmid DNA (Amersham Biosciences). The PCR-based overlap extension method (12) was used to produce an active-site knockout mutant of SARS CoV M^{pro} with His-41 replaced by Ala (H41A) using pGEX-6p-1–SARS-CoV M^{pro} as a template. The primers were designed so that the ends of the two PCR products contained complementary sequences, which allowed the two fragments to be spliced in a second PCR. The four primers used for the single point mutation were the following: 5′-CGGGATCCAGTGGTTTTAGG AAAATG-3′ (forward A), 5′-CCGCTCGAGTCATTGGAAGGTAACACCAGA-3′ (reverse A), 5′-A ATGACCGCTCTTGGACAGTATACTGT-3′ (forward B), and 5′-CCAAGAG CGGTCATTTGCACAGCAGAA-3′ (reverse B). Specifically, in the first PCR two sets of primers (forward A/reverse B and forward B/reverse A) were used to generate the templates for the second PCR. The two primers (forward A/reverse A) were used in the second PCR, and then the PCR products were inserted into the BamHI and XhoI sites of the pGEX-6p-1 plasmid. The resulting plasmids containing the H41A mutation were verified by sequencing and then transformed into *E. coli* BL21 (DE3) cells. The protein expression and purification of the SARS-CoV M^{pro} were described previously (35). The crystallization of SARS-CoV M^{pro} (H41A) was the same as that for the wild-type protease (33, 35). An 11-amino-acid peptidyl substrate of the sequence TSAVLQSGFRK was dissolved at a 20 mM concentration in 7.5% (wt/vol) PEG6K, 6% (vol/vol) dimethylsulfoxide (DMSO), and 0.1 M morpholineethanesulfonic acid (Mes) (pH 6.0). A 3-μl aliquot of this solution was added to the crystallization drop (3 μl), and the crystals were soaked for 8 days before data collection.

Diffraction data collection. A total of four data sets were collected (Table 1). Data for the SeMet IBV M^{pro} derivative were collected to a 2.8-Å resolution at the peak wavelength (for the maximum f'') at 100°K using a Structural Biology Center (2,000 by 2,000) charge-coupled display detector on beamline BL19-ID of the Advanced Photon Source, Argonne National Laboratory. The cryoprotectant solution contained 20% (vol/vol) glycerol, 2% (wt/vol) PEG4K, 9.6% (vol/vol) 2-propanol, and 0.08 M sodium cacodylate (pH 6.5). Another data set for the native IBV M^{pro} was collected to a 2.35-Å resolution at 100°K on beamline BL-5A at Photon Factory (KEK, Japan) using an ADSC Q315 e-coupled display detector. Data for the IBV M^{pro}-N3 complex and SARS-CoV M^{pro} H41A mutant peptidyl substrate complex were collected at 100°K in house with a Rigaku CuKα rotating-anode X-ray generator (MM007) at 40 kV and 20 mA (1.5418 Å) and using a Rigaku R-AXIS IV++ image plate detector. The IBV M^{pro} complex crystal was used directly in data collection without a cryoprotectant. The cryoprotectant solution for the SARS-CoV M^{pro} mutant complex contained 30% PEG400 and 0.1 M Mes (pH 6.0). All data integrations and scaling were performed using HKL2000 (23). The Matthews coefficient of the new IBV M^{pro} crystal form suggested the existence of three protein molecules per asymmetric unit with an estimated solvent content of 54%.

Structure solution, refinement, and analysis. The IBV M^{pro} structure was solved by the single-wavelength anomalous dispersion method (11) using the diffraction data set collected at the peak wavelength for selenium. The analysis of the selenium positions, performed with the program SHELXD (27), located all 12 expected selenium sites (four in each protein molecule). Phasing and density modifications subsequently were performed with SOLVE (32) and RESOLVE (31). The resulting electron density maps were of sufficient quality for chain tracing. Molecular replacement performed with CNS (4) was employed for tracing the typical homodimer (named molecules A and B) into the electron density map using the crystal structure of human CoV-229E (HCoV-229E) M^{pro} as a starting model (Protein Data Bank code 1P9S). The third M^{pro} molecule (named molecule C) was clearly identified in the electron density map, and its tracing was facilitated using the noncrystallographic symmetry of the selenium positions. Cycles of manual adjustment to the model with Coot (8) and subsequent refinement using REFMAC (21) led to a final model with a crystallographic *R* factor (R_{cryst}) of 22.7% and a free *R* factor (R_{free}) of 25.9% at 2.35-Å resolution.

The IBV M^{pro}-N3 complex structure was determined by the molecular replacement method implemented in CNS using the homodimer (molecules A and B) from the above-described native IBV M^{pro} structure as the search model. Manual adjustments to the model were made with the program O (22), and subsequent refinement was performed in CNS. Data quality and refinement statistics are summarized in Table 1.

The structure of the mutant protein (SARS-CoV M^{pro} H41A) in complex with its N-terminal peptide substrate was determined by the molecular replacement method using a SARS-CoV M^{pro} monomer (Protein Data Bank code 1UK2) (35) as a search model. In the complex structure, there are two M^{pro} molecules (named A and B) per asymmetric unit, and it forms a symmetrical homodimer. An 11-mer peptide was identified in molecule A and an 8-mer peptide in molecule B from the initial difference electron density maps. The validation of all final models was carried out with PROCHECK (17).

In vitro inhibition assays. Proteolytic activity assays of IBV M^{pro} have been described previously (33, 34). The fluorogenic substrate of SARS-CoV M^{pro}, MCA-AVLQSGFR-Lys(Dnp)-Lys-NH2 (>95% purity; GL Biochem Shanghai Ltd., Shanghai, China), was used to assess the activity of IBV M^{pro}. The excitation and emission wavelengths of the fluorogenic substrate were 320 and 405 nm, respectively. The assay was performed in a buffer of 50 mM Tris-HCl (pH 7.3) and 1 mM EDTA at 30°C, and kinetic parameters were determined by following our previous work (34).

In ovo inhibition. Titers of the IBV M41 viruses were established as follows. The virus was serially 10-fold diluted in phosphate-buffered saline (PBS) and then inoculated into the allantoic cavity of 10-day-old specific-pathogen-free (SPF) chicken embryos (six embryos per dilution and 0.1 ml virus dilution per embryo). The embryos were incubated at 37°C and were inspected daily. Eight days after inoculation, the eggs were opened and examined to check for typical lesions (including crispature and dwarfism in embryos, yolk sac shrinking, an increase in allantoic fluid, and lithate deposits on the midkidney of embryos) that might signify IBV infection. Six embryos inoculated with PBS were used as negative controls, and another six uninoculated embryos were used as blank controls. The dilution that could cause 50% of embryos to be infected by IBV was calculated using the method described by Reed and Muench (26) and determined as the virus titer (50% egg infectious dose [EID_{50}]).

To assess whether N3 could be used as an anti-IBV preventive agent or a curative agent, two groups of in ovo inhibition experiments were performed. For the curative group, a series of doses of N3 (0.02 to 0.64 μmol) was injected into the allantoic cavity of 10-day-old SPF chicken embryos 3 h (for eight embryos; repeated per dose of N3) or 6 h (for six embryos; repeated per dose of N3) after inoculation by a 100-EID_{50} titer of IBV M41 virus. For the preventive group, N3 was preinjected into the embryos 3 h (for eight embryos; repeated per dose of N3) or 6 h (six embryos; repeated per dose of N3) prior to the inoculation by a 100-EID_{50} titer of virus. Eight days after inoculation, the eggs were opened to check if the embryos were infected by IBV. The inhibitor dose that could protect 50% of embryos from IBV infection was calculated using the method described by Reed and Muench (26) and expressed as the 50% protective dose (PD_{50}).

Meanwhile, a preliminary toxicity assay was performed to assess any potential adverse effects of N3 on the development of chicken embryos. The highest dose of N3 (0.64 μmol) dissolved in DMSO was injected into 16 embryos. Sixteen embryos inoculated with DMSO were used as negative controls, while another 16 uninoculated embryos were used as blank controls. Eight days after inoculation, half of the eggs were opened and examined for pathological changes to the organs of the embryos. The remainder of the eggs were continuously incubated at 37°C until the chickens were hatched. All in ovo experiments were performed in a biosafety level 2 bioprotective laboratory.

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

TABLE 1. Data collection and refinement statistics

Parameter	Data set for:			
	Se-Met IBV M$^{\text{pro}}$	Native IBV M$^{\text{pro}}$	IBV M$^{\text{pro}}$ N3	SARS-CoV M$^{\text{pro}}$ H41A substrate
Data collection statistics				
Wavelength (Å)	0.9795	1.0000	1.5418	1.5418
Resolution (Å)	50–2.8 (2.91–2.80)^{b}	50–2.35 (2.43–2.35)	50–2.00 (2.07–2.00)	50–2.40 (2.49–2.40)
Space group	P6$_1$22	P6$_1$22	P1	P2$_1$
Cell parameters				
a (Å)	118.2	118.9	53.2	52.0
b (Å)	118.2	118.9	54.5	95.8
c (Å)	267.7	270.9	66.7	67.7
α (°)	90.0	90.0	111.1	90.0
β (°)	90.0	90.0	104.3	102.9
γ (°)	120.0	120.0	91.3	90.0
Total reflection	713,639	339,766	165,955	82,777
Unique reflection	56,512	47,480	42,883	25,190
Completeness (%)	100.0 (100.0)	98.9 (99.8)	94.2 (82.6)	99.8 (99.9)
Redundancy	12.6 (8.6)	7.2 (7.3)	3.9 (3.3)	3.3 (3.3)
$R_{\text{merge}}{}^{a}$	0.170 (0.715)	0.054 (0.358)	0.041 (0.225)	0.106 (0.474)
Sigma cutoff	0	0	0	0
I/σ (I)	16.6 (2.5)	39.8 (5.3)	30.4 (5.1)	11.8 (2.5)
Refinement statistics				
Resolution range (Å)		50–2.35	50–2.00	30–2.50
$R_{\text{work}}{}^{c}$ (%)		22.7	21.6	19.9
R_{free} (%)		25.9	24.2	26.7
RMSD from ideal geometry				
Bonds (Å)		0.009	0.011	0.007
Angles (°)		1.62	1.75	1.39
Average B factor (Å^2)				
Main chain		50.3	40.4	29.7
Solvent		56.4	49.9	42.1
Ramachandran plot^{d}				
Favored (%)		85.7	91.6	84.3
Allowed (%)		14.0	8.4	14.4
Generously allowed (%)		0.3	0.0	0.7
Disallowed (%)		0.0	0.0	0.6

[a] $R_{\text{merge}} = \Sigma I_i - \langle I \rangle / \Sigma I$, where I_i is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity of that reflection.

[b] $R_{\text{work}} = \Sigma F_p - F_c / \Sigma F_p$, where F_c is the calculated and F_p is the observed structure factor amplitude.

[c] Ramachandran plots were generated by using the program PROCHECK.

[d] Numbers in parentheses correspond to the highest-resolution shell.

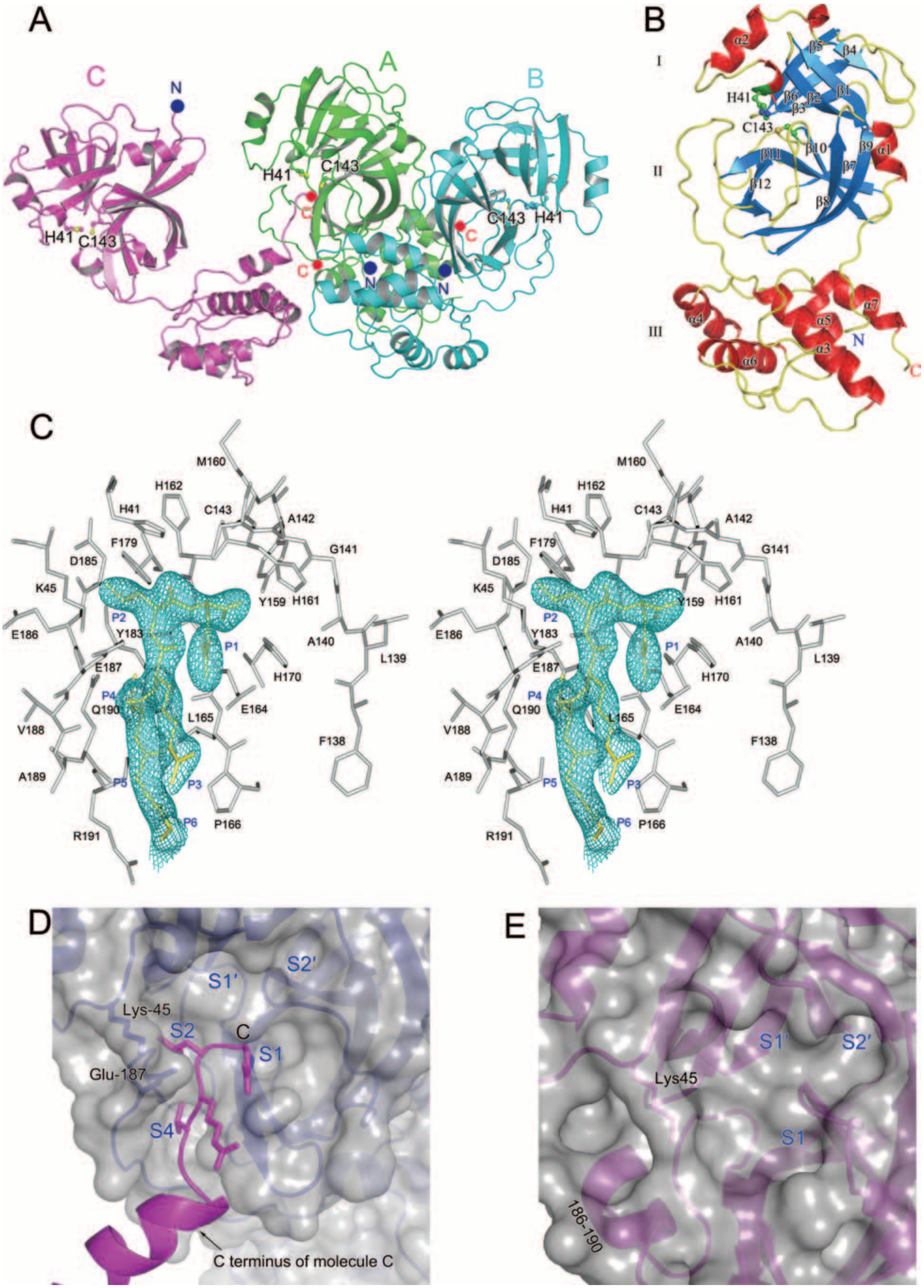
Accession codes. Coordinates and structure factors for IBV M$^{\text{pro}}$, IBV M$^{\text{pro}}$ in complex with inhibitor N3, and the SARS-CoV M$^{\text{pro}}$ H41A mutant in complex with an N-terminal substrate have been deposited in the Protein Data Bank under accession numbers 2Q6D, 2Q6F, and 2Q6G, respectively.

RESULTS

Overall structure of native IBV M$^{\text{pro}}$. The IBV M$^{\text{pro}}$ crystal structure at a 2.35-Å resolution shows three M$^{\text{pro}}$ molecules, named A, B, and C, per asymmetric unit (Fig. 1A), which is unique among all CoV M$^{\text{pro}}$ structures reported to date. While molecules A and B form a typical catalytically active and symmetrical homodimer, molecule C is not involved in such a dimer. Instead, its C terminus inserts into the substrate binding site of molecule A (Fig. 1A). Molecules A and B are quite similar, with an RMSD (root mean square deviation) of 1.1 Å for all equivalent Cα atoms, while molecule C is less similar to either A or B, having a mean RMSD of 2.5 Å for the Cα atoms of residues 6 to 183.

Each IBV M$^{\text{pro}}$ molecule is comprised of three domains, I to III (Fig. 1B). Domains I and II (i.e., residues 3 to 99 and 100 to 182, respectively) have a chymotrypsin-like, two-β-barrel fold in common with the M$^{\text{pro}}$ structures of transmissible gastroenteritis CoV (TGEV), HCoV-229E, and SARS-CoV (1, 2, 35). Domain III (residues 199 to 307) of IBV M$^{\text{pro}}$ consists of five α helices that adopt a globular structure apparently unique to CoV M$^{\text{pro}}$. Domains II and III are connected by a loop of residues 183 to 198, which exhibits two distinct conformations in the three M$^{\text{pro}}$ molecules. In molecules A and B, it assumes a fairly extended conformation; in molecule C, however, residues 186 to 190 form a short helix (Fig. 1E). The substrate binding sites are located in the deep cleft between domains I and II, with the catalytic dyad formed by His-41 and Cys-143 at the center of this cleft. Each subunit contains one substrate binding site contributed mainly from itself. Nevertheless, the two monomers swap their N termini to stabilize the S1 pocket in the IBV M$^{\text{pro}}$ dimer; similar swapping was also observed in the M$^{\text{pro}}$ structures of TGEV, HCoV-229E, and SARS-CoV (1, 2, 19, 33, 35). This arrangement may explain the require-

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV



Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

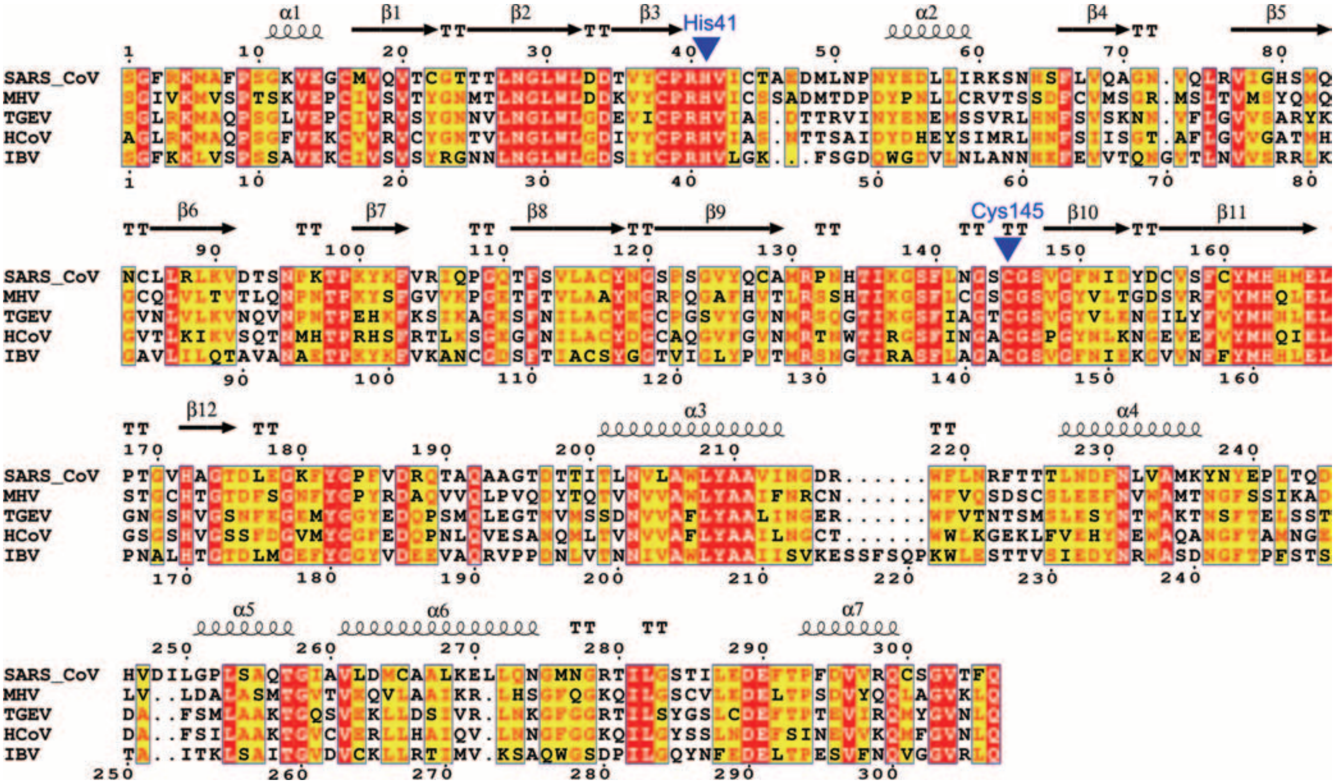


FIG. 2. Structure-based sequence alignment of the main proteases of CoV from all three groups. SARS-CoV, SARS-CoV (group II); MHV, mouse hepatitis virus (group II); TGEV, porcine TGEV (group I); HCoV, HCoV 229E (group I); and IBV, avian IBV (group III). Secondary structures of SARS-CoV M^{pro} are indicated above the sequence. Residue numbers of SARS-CoV M^{pro} (above) and IBV M^{pro} (below) also are indicated. The catalytic dyad His-41 and Cys-145 (SARS-CoV M^{pro}) are labeled.

ment of dimerization for the full activity of the M^{pro} proteins (1, 2, 19, 33, 35).

According to a structure-based sequence alignment (Fig. 2), there is one deletion and two insertions in IBV M^{pro} not found in the M^{pro} of TGEV and HCoV-229E. The two insertions, namely, residue 70 and residues 216 to 221, all are located in loop regions with unknown functional significance. The three-residue deletion after Leu-50 makes the corresponding loop (i.e., residues 44 to 50) much tighter than the equivalent region in TGEV and HCoV-229E M^{pro}. The side chain of Lys-45 in this loop is involved in the formation of the S2 pocket, corresponding to Thr-47 in TGEV M^{pro} and HCoV M^{pro} and Met-49 in SARS-CoV M^{pro}. Therefore, the S2 subsite appears to be unique in IBV M^{pro}.

Substrate binding sites of IBV M^{pro}. In the IBV M^{pro} structure, the substrate binding pockets of molecule A are occupied by the C terminus (residues 302 to 307, corresponding to the P6 to P1 sites of the M^{pro} substrate) of molecule C (Fig. 1C and D), which forms an antiparallel β sheet with β11 (residues 163 to 166) in domain II and with residues 188 to 190 of the linker loop between domains II and III.

In this A-C complex, the S1, S2, and S4 substrate binding sites of molecule A can be clearly recognized (Fig. 1D). The side chains of Phe-A138, His-A161, Glu-A164, and His-A170 are involved in constituting the S1 subsite, which has an absolute requirement for Gln at the P1 position via two hydrogen bonds (1, 2, 35). Nevertheless, the side chain of Gln-307 of molecule C does not fit well into the S1 pocket. Instead, its side chain is flipped out from the pocket, probably because the availability of the main chain carboxyl group of Gln in this case (the distance between the carboxyl carbon of Gln-C307 and the sulfur atom of the catalytic Cys-A143 is ~3.1 Å). As a result,

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

FIG. 1. Three-dimensional structure of IBV M^{pro}. (A) Overall structure of IBV M^{pro} in one asymmetric unit. Molecules A (green) and B (cyan) form a homodimer, with the C terminus of molecule C (magenta) inserted into the substrate binding pocket of molecule A. Catalytic dyads are indicated, and the N and C termini are labeled by blue and red spheres and the letters N and C, respectively. (B) Subunit of IBV M^{pro} (molecule B). α Helices are colored red, β strands are colored blue, and loops are colored yellow. Domains I, II, and III and the catalytic dyad residues His-41 and Cys-143 are indicated. (C) A stereo view showing the C terminus of molecule C bound into the substrate binding site of molecule A. The C302 to C307 residues are shown in gold and are covered by an omit map at 2.35-Å resolution contoured at 1.2 σ. Residues forming the substrate binding pocket in molecule A are shown in silver. (D) Surface of the substrate binding sites of molecule A in the IBV M^{pro} structure. The S1′, S2′, S1, S2, and S4 subsites are labeled, and the C terminus of molecule C, which occupies the substrate binding sites, is colored magenta. (E) Surface of the substrate binding sites of molecule C in IBV M^{pro}. The S1′, S2′, and S1 subsites are labeled, and residues 186 to 190, which form a novel helix, also are labeled.

the side chain of Gln-P1 is more or less flexible and forms only one hydrogen bond with Glu-A164. The oxyanion hole is not properly formed by Gly-A141, Ala-A142, and Cys-A143, probably due to a disturbance by the flexible Gln, in contrast to the correctly folded oxyanion holes in molecule B. The side chains of His-A41, Lys-A45, Leu-A163, Phe-A179, Asp-A185, and Glu-A187 are involved in forming the deep hydrophobic S2 subsite that is able to accommodate the relatively large side chain of conserved Leu or, in a few cases, Val in the substrates of IBV M^{pro}. As expected, the side chain of Leu-P2 (Leu-C306) is well oriented into the S2 hydrophobic pocket and stabilized by van der Waals interactions. The side chain of Arg-P3 (Arg-C305) is oriented toward bulk solvent but also interacts with the side chain of Glu-164 via van der Waals interactions. The side chains of Leu-A163, Leu-A165, Tyr-A183, and Gln-A190 form the relatively small hydrophobic S4 subsite, which should be able to accommodate small residues such as Val, Ser, Thr, Ala, or Pro. The Gly-P5 and Gly-P6 residues are in faint interactions with the protease. No other interaction is observed between molecule C and molecules A and B from the same asymmetric unit in the IBV M^{pro} crystal structure.

The monomeric form of IBV M^{pro}. Molecule C presents a novel conformation distinct from those of the other two M^{pro} molecules in the IBV M^{pro} structure. The superposition of the first two domains in molecules C and A confirmed that they share similar domain structures (Fig. 3). However, they bear clear structural differences at the whole-molecule level, mostly due to the conformational change in the linker region connecting the N-terminal two-β-barrel domains (domains I and II) with the C-terminal α-helical domain III. This conformational change includes the formation of a short helix (residues 186 to 190) in this linker region (Fig. 3), which results in a nearly 5-Å movement of domain III away from domains I and II in molecule C. Differences also occur in the N- and C-terminal conformations between molecules C and A. As described above, the C terminus of molecule C fits well into the substrate binding pocket of molecule A, which was not observed in those of molecules A and B. At the other end, the N terminus of molecule C is flexible and directed away from the surface of domain I; thus, residues 1 to 5 in molecule C could not be traced in the electron density map. In contrast, the N terminus of molecule A inserts into the dimer interface formed by its own domains II and III as well as domain II of the neighboring subunit, where it makes a number of specific interactions to stabilize the dimer structure. This monomer structure of IBV M^{pro} reveals a significant structural flexibility of the linker region connecting domains II and III that has not been reported for other structures of dimeric CoV M^{pro}s to date. The presence of the monomeric form probably was triggered by the binding and fixation of its C terminus in the active site of the M^{pro} dimer, which may preclude dimerization.

In the absence of dimerization, the substrate binding sites of molecule C are not well organized. Only the S1, S1′, and S2′ subsites maintain their correct conformations (Fig. 1E). The S2 and S4 subsites collapse, partly because residues 186 to 190 in the linker region adopt an unusual helical conformation (Fig. 1E). Nevertheless, the flexibility in the linker region may allow incidental activity in molecule C in the absence of dimerization, which is required for the maturation of M^{pro}. In contrast, in the homodimer form the linker region adopts a conformation to achieve the highest level of proteolytic activity.

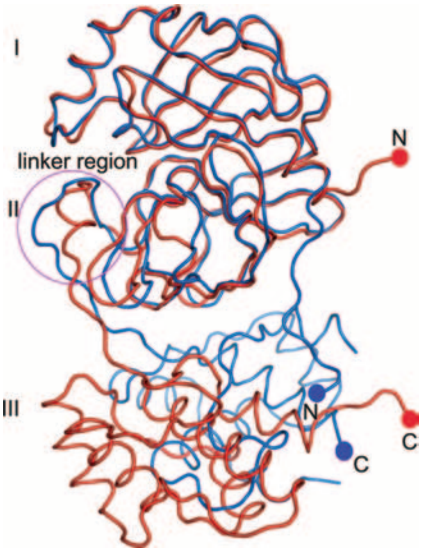


FIG. 3. Superposition of the first two domains in molecules C (red) and A (blue) of the IBV M^{pro} structure. The structures of domains I and II are quite similar. While domains III from the two proteins also are quite similar (with a Cα RMSD of 0.5 Å), its location in molecule C is transformed away from domains I and II by a conformational change in the long linker region (labeled in the figure) connecting domains II and III.

Overall structure of SARS-CoV M^{pro} H41A mutant in complex with its N-terminal substrate. To further investigate the substrate binding and specificity of CoV M^{pro}, we crystallized an active-site knockout mutant, H41A, of SARS-CoV M^{pro}, soaked the crystals with its natural, N-terminal peptide substrate, and determined the complex crystal structure at a 2.5-Å resolution. There are two M^{pro} molecules per asymmetric unit in this complex structure, named A and B, which form a typical M^{pro} dimer. Both subunits have the same overall structure and almost identical substrate binding modes. An 11-amino-acid peptide in subunit A and an 8-amino-acid peptide in subunit B were identified from difference Fourier electron density maps. The enzyme-bound 11-mer peptidyl substrate essentially is comprised of two parts, the N-terminal residues P6 to P1 and the C-terminal residues P1′ to P5′, which roughly assume conformations of two separate β strands (Fig. 4A). Similarly to the conformation of the C-terminal residues observed in the IBV M^{pro} crystal structure, residues P6 to P1 form an antiparallel β sheet with residues 164 to 168 on one side and residues 189 to 191 of the linker loop between domains II and III on the other side (Fig. 4B). The P1′ to P5′ strand is located in a groove formed by β2 (residues 24 to 27) and the loop of residues 142 to 144 near the catalytic Cys-145 (Fig. 4A and B).

Substrate binding sites of SARS-CoV M^{pro}. On the N-terminal side of the substrate, the P6 to P1 positions (Thr-Ser-Ala-Val-Leu-Gln) share a similar binding mode with the previously reported SARS-CoV M^{pro} structures in complex with a variety of Michael acceptor inhibitors (34). In particular, in the S1 subsite the Gln residue required for high cleavage efficiency seems to intercalate more naturally than the lactam ring in the Michael acceptor inhibitors that we previously designed (34).

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

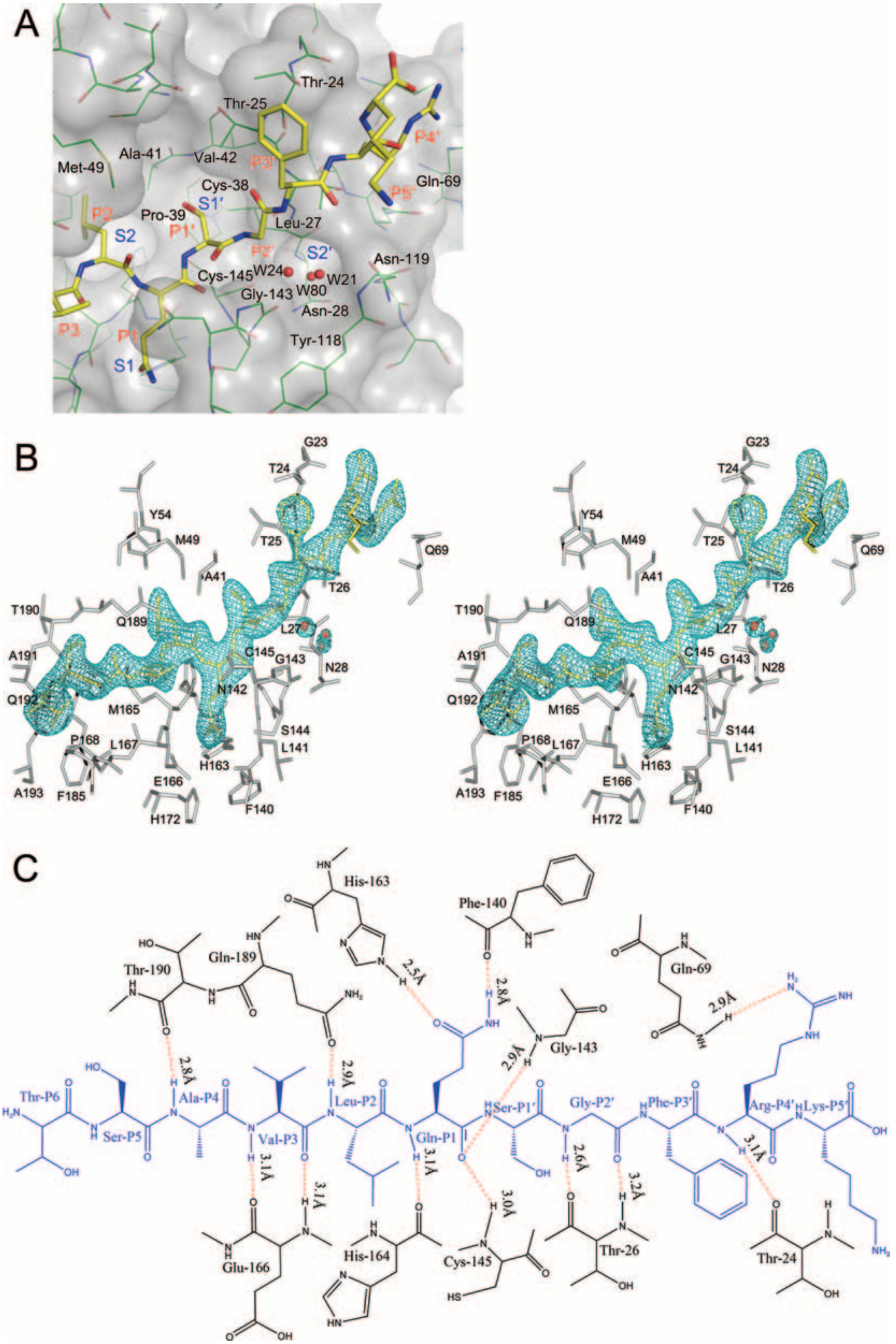


FIG. 4. Structure of the SARS-CoV M^{pro} H41A mutant in complex with an N-terminal 11-peptidyl substrate. (A) Surface representation of SARS-CoV M^{pro} H41A mutant (white) in complex with the N-terminal substrate (yellow). Positions of P3 to P5′, S1 to S2′, and residues forming the S1′, S2′ sites are labeled. Notice that there are three water molecules (shown as red spheres) occupying the S2′ pocket. (B) Stereo view showing the N-terminal peptide substrate bound into the substrate binding pocket of the SARS-CoV M^{pro} H41A mutant. The substrate is shown in gold and is covered by an omit map at 2.5-Å resolution contoured at 1.2 σ. Residues forming the substrate binding pocket are shown in silver. Three water molecules (in red) occupy the S2′ pocket. (C) Schematic diagram of the interactions between the N-terminal 11-peptidyl substrate and the SARS-CoV M^{pro} H41A mutant. The substrate is shown in blue. Hydrogen bonds are shown as dashed lines, and interaction distances are marked.

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

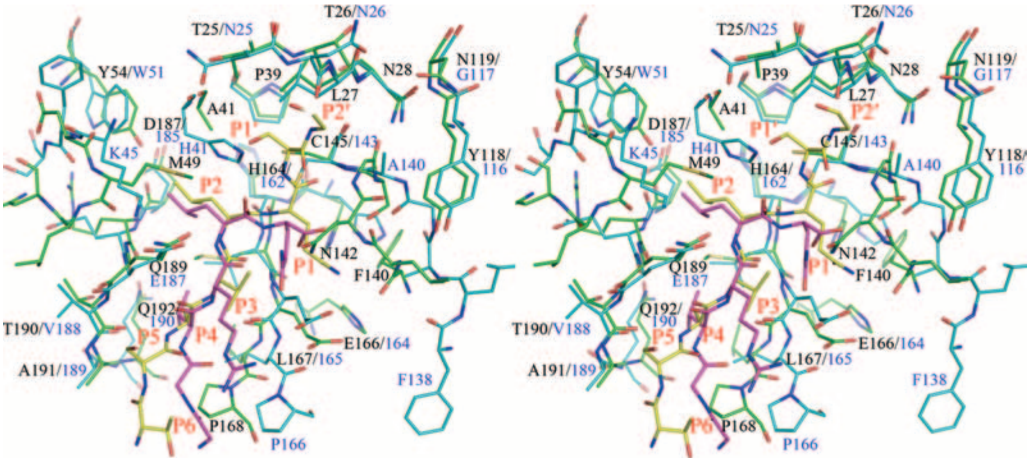


FIG. 5. Superposition of the substrate-binding pockets of IBV M^{pro} (molecule A) and SARS-CoV M^{pro} mutant-substrate complex (in stereo). The C terminus of molecule C (P6 to P1 sites) in IBV M^{pro} (cyan) is in magenta, and the peptidyl substrate of SARS-CoV M^{pro} (green) is in yellow. Residues of SARS-CoV M^{pro} are labeled in black, and residues of IBV M^{pro} are labeled in blue.

Two strong hydrogen bonds, between the Oε1 atom of Gln-P1 and the Nε2 atom of His-163 (2.5 Å) and between the Nε2 atom of Gln-P1 and the main chain carbonyl oxygen of Phe-140 (2.8 Å), ensure that the conserved Gln-P1 residue comfortably fits in the S1 pocket (Fig. 4C). The latter hydrogen bond has not been reported for previous enzyme-inhibitor complex structures. The carbonyl oxygen of Gln-P1 is stabilized by the oxyanion hole formed by the amide groups of Gly-143 and Cys-145 (Fig. 4C). The P2 to P4 residues bind to the enzyme similarly to the previously reported peptidyl inhibitors (34). In addition, the Ser-P5 and Thr-P6 residues interact with Pro-168 and Ala-191 of the enzyme through van der Waals interactions.

On the C-terminal side, no structural information for the binding mode of P1′ to P5′ residues with M^{pro} has previously been reported. Therefore, the complex structure presented here allows us to explore the substrate binding and specificity of S1′ to S5′ in SARS-CoV M^{pro} in an unprecedented way (Fig. 4). Small residues such as Ser, Gly, and Ala are preferred at the relatively shallow S1′ subsite, which is composed of Thr-25, Leu-27, Cys-38, Pro-39, Ala-41, Val-42, and Cys-145. The small P1′ residue directly interacts with the side chains of Thr-25, Leu-27, and Cys-145 via van der Waals interactions. The S2′ subsite is a narrow but deep pocket composed of residues Thr-26, Asn-28, Tyr-118, Asn-119, and Gly-143. In our complex structure, the S2′ subsite is occupied by Gly-P2′, with additional space occupied by three ordered water molecules (W21, W24, and W80). The hydrophilic S2′ pocket can accommodate a long side chain residue at the P2′ position, such as the lysine residue at the corresponding site for its C-terminal autocleavage. The main-chain amide and the carbonyl oxygen of Gly-P2′ form a pair of hydrogen bonds with the main-chain atoms of Thr-26 (Fig. 4C). The P3′ side chain appears to point toward the solvent and makes no specific interactions with the protease. The Arg-P4′ residue also is stabilized by two hydrogen bonds: one occupies 3.1 Å between the amide group of Arg-P4′ and the carbonyl oxygen of Thr-24, and the other occupies 2.9 Å between the Nη1 atom of Arg-P4′ and the Nε2 atom of Gln-69. The complex structure shows that the P5′ residue has little interaction with the protease.

Active-site comparison between IBV M^{pro} and SARS-CoV M^{pro}. Since the substrate-bound structures of both IBV M^{pro} and SARS-CoV M^{pro} became available from this study, we compared the conformations of the active sites in these two structures (Fig. 5). In the S1 subsite, the outer wall made up of residues 141 to 143 in the SARS-CoV M^{pro} structure is not present in the IBV M^{pro} structure, possibly due to the disturbance of Gln-P1 (Gln-C307). Ala-140 of IBV M^{pro} is away from the active site, so that the S1 pocket is larger than that in SARS-CoV M^{pro}. Lys-45 and Glu-187 in IBV M^{pro}, instead of Met-49 and Glu-189 in SARS-CoV M^{pro}, form the outer wall of the S2 subsite (Fig. 1D). Lys-45 of IBV M^{pro} moves ~2 Å away from the S2 subsite, such that the S2 pocket in IBV M^{pro} is slightly larger than that in SARS-CoV M^{pro}. The P3 position of IBV M^{pro} is occupied by an arginine residue with a long side chain, which makes interactions with the side chains of Glu-164 and Arg-B305. It seems likely that a longer side chain is preferred to stabilize the substrate binding site here and that the modification of the P3 position may be a good choice for the design of substrate-based inhibitors targeting CoV M^{pro}. The S1′ and S2′ subsites are quite similar in both M^{pro} structures, implying substrate conservation on the two subsites, which also may be applicable for inhibitor design.

Structure of IBV M^{pro} in complex with inhibitor N3. We have previously designed a series of broad-spectrum inhibitors targeting CoV M^{pro} (34). Of these inhibitors, a Michael acceptor inhibitor named N3 strongly inhibits the replication of SARS-CoV, TGEV, HCoV-229E, mouse hepatitis virus A59, and feline infectious peritonitis virus in cell-based assays (34). In this study, the cocrystallization of N3 with IBV M^{pro} yielded high-quality crystals. The subsequent high-resolution structure of IBV M^{pro} in complex with N3 together with the in vitro inhibition assay results (shown in Table 2) reveal that N3 could block the activity of the M^{pro} through a standard Michael addition reaction.

Unlike the native structure, the complex structure of IBV

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

TABLE 2. Enzyme activity and enzyme inhibition data for IBV M^{pro} and SARS-CoV M^{pro}

Enzyme and inhibitor	K_m (μM)	k_{cat} (s^{-1})	K_i (μM)	k_3[a] (10^{-3} s^{-1})	k_3/K_i (M^{-1} s^{-1})
IBV M^{pro}	119 ± 14	0.23 ± 0.02			
N3			3.6 ± 0.4	25.3 ± 1.4	(7.1 ± 0.6) × 10^3
N27			2.6 ± 0.3	22.9 ± 2.1	(8.7 ± 0.4) × 10^3
H16			2.8 ± 0.2	21.0 ± 1.4	(7.5 ± 0.5) × 10^3
SARS-CoV M^{pro}	40.0 ± 0.8	1.06 ± 0.04			
N3			9.6 ± 1.0	142 ± 28	(15.0 ± 2.8) × 10^3
N27			3.1 ± 0.2	61.3 ± 4.6	(20.0 ± 0.7) × 10^3
H16			3.3 ± 0.5	89 ± 20	(27.0 ± 4.8) × 10^3

[a] k_3, activation rate constant for covalent bond formation.

M^{pro} with the inhibitor N3 has a homodimer in each asymmetric unit. Each dimer has approximate C_2 symmetry, which is consistent with other M^{pro}-inhibitor complex structures we have solved to date (33, 34). From the omit electron density map, clear electron density was identified for N3 bound in the substrate binding pocket (Fig. 6B and C). Residues P3 to P5 form a typical antiparallel β sheet with residues 163 to 166 of the β11 strand on one side, while on the other side they interact with residues 187 to 189 of the loop linking domains II and III.

In the inhibitor-bound complex structure, the S_γ atom of the nucleophilic Cys-143 forms a clear 1.9-Å C-S covalent bond with the C_β atom of the vinyl group, which is a typical Michael addition (Fig. 6B). The fact that Michael acceptor inhibitors can irreversibly react with the active site of the enzyme makes N3 a standard suicide inhibitor. We have previously reported the crystal structure of SARS-CoV M^{pro} in complex with N3 (33, 34), so we superimposed the substrate binding pockets of the IBV M^{pro}-N3 and SARS-CoV M^{pro}-N3 complex structures. A comparison of the two inhibitor-bound complex structures implies a similar binding mode of this Michael acceptor inhibitor (Fig. 6D). The largest difference between the two complex structures occurs, however, in the orientation of the benzyl ester group. The side chain of Asn-142 in the SARS-CoV M^{pro} complex structure disturbs the comfortable orientation of the benzyl ester at the P1′ site. In the corresponding position of IBV M^{pro}, Asn-142 is replaced with Ala-140, and the benzyl ester points toward the solvent in a much more comfortable orientation. Another significant difference lies in the S2 site, where Lys-45 in the IBV M^{pro} complex structure is replaced by Met-49 in the SARS-CoV M^{pro} complex structure.

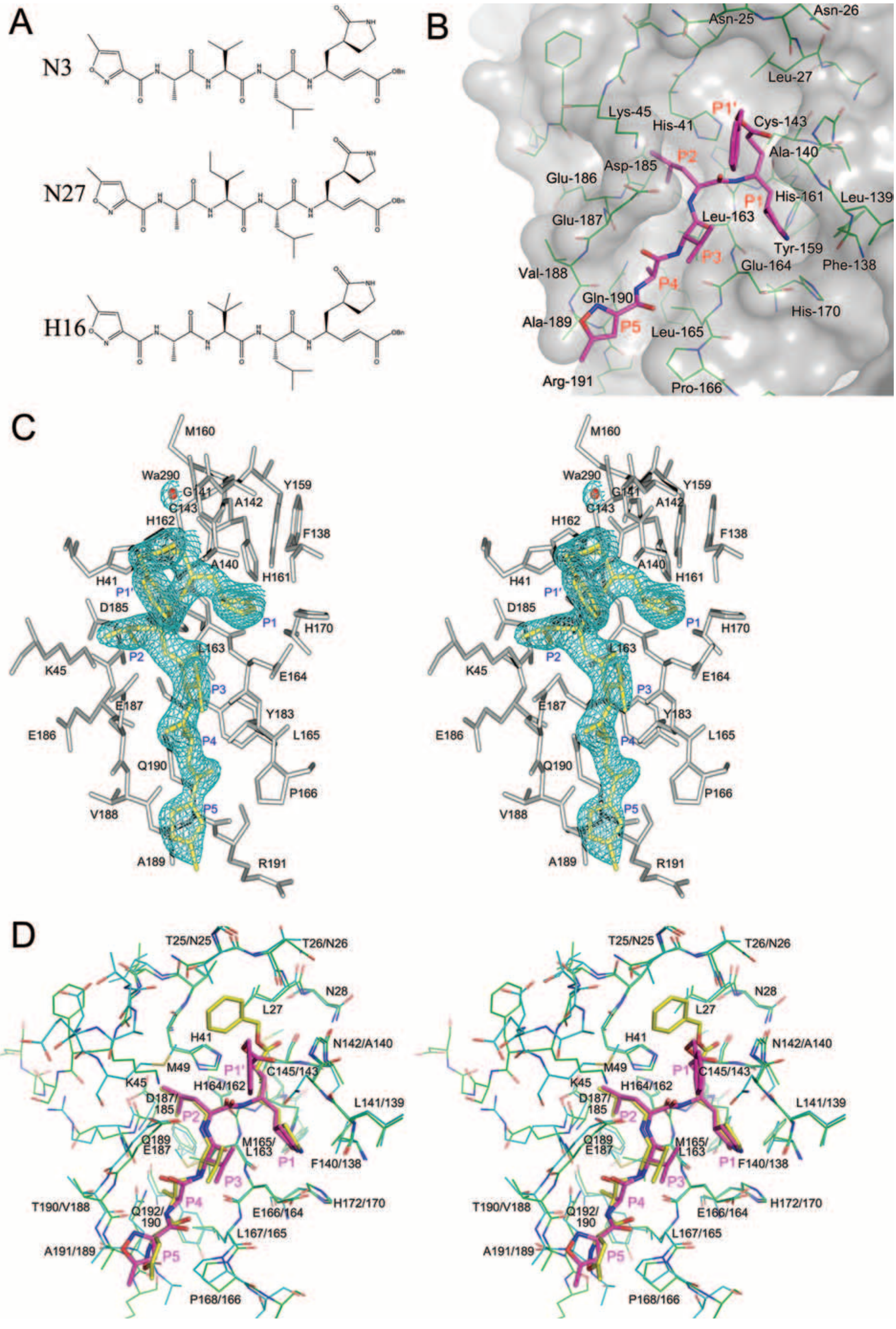
In ovo inhibition of IBV by N3. An in ovo inhibition assay in chicken embryos was performed to further substantiate the effects of N3 on IBV inhibition. One method used was the neutralization test in chicken embryos, which was implemented to assess the neutralizing power of an antiserum or inhibitor against pathogens such as viruses (6). Infection by the IBV M41 strain was identified by the presence of typical lesions (as described in Materials and Methods). Firstly, the virus titer (EID$_{50}$) of this IBV M41 strain was determined as 0.1 ml, a $10^{-6.5}$ dilution of viruses. To assess the stage of infection at which the inhibitor can be used effectively, a series of doses of N3 was used as curative agents and introduced into the chicken embryos 3 h (Fig. 7A) or 6 h (Fig. 7B) following inoculation with a 100-EID$_{50}$ titer of IBV M41 virus. The dose-response data show that N3 is able to penetrate cells to inhibit the replication of IBV viruses, probably at the beginning of infection (Fig. 7A and B). The PD$_{50}$ of N3 was calculated as 0.13 μmol for the 3-h group and 0.17 μmol for the 6-h group according to the method described by Reed and Muench (26). These inhibition data further imply that the earlier N3 is used during infection, the more effective is the inhibition of the IBV virus. For instance, a 0.08-μmol dose of N3 per embryo introduced 3 h after inoculation could protect ~40% of chicken embryos not infected by the IBV M41 virus, while it could protect no chicken embryos when introduced 6 h after inoculation. However, a 0.64-μmol dose of N3 per embryo introduced either 3 or 6 h after inoculation could protect, in both cases, all chicken embryos from infection.

To verify whether N3 could be used as an anti-IBV preventive agent, another group of experiments was performed. A series of doses of N3 was introduced into the chicken embryos 3 h (Fig. 7C) or 6 h (Fig. 7D) prior to virus inoculation. The PD$_{50}$ of N3 for this preventative group was calculated as 0.099 μmol for the 3-h group and 0.095 μmol for the 6-h group. Therefore, consistently with the antiviral activity of N3 in vitro, our results show better inhibition of IBV with N3 used as a preventive agent than as a curative agent.

Meanwhile, in the preliminary toxicity assay of N3 on chicken embryos, half of the eggs from the three groups (0.64-μmol N3 control group, DMSO-negative control group, and blank control group) were opened and examined 8 days after inoculation. No significant pathological changes or lesions were found in the shape and organs of chicken embryos. The remaining eggs from the three groups then were hatched, and no significant differences were found in the baby chickens physically or spiritually. Thus, the toxicity assay for N3 signifies that even a 0.64-μmol dose of N3 per embryo has no detectable negative impact on the development of the chicken embryos.

Further improvements on inhibitor design. The substrate-bound structures of IBV M^{pro} and SARS-CoV M^{pro} provide useful information for antiviral drug design. Taking the framework of N3 as a starting point, we designed another series of Michael acceptor inhibitors and measured their inhibition activities against SARS-CoV M^{pro} and IBV M^{pro}. Two of these new inhibitors, named N27 and H16 (Fig. 6A), which were synthesized by the State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, China, have a relatively large side chain at the P3 position and show

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV



Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

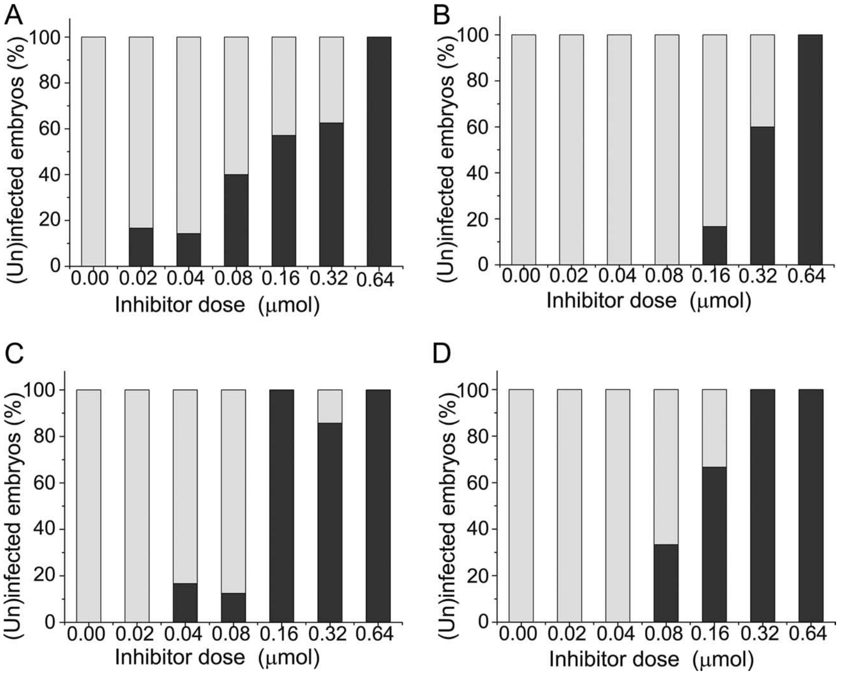


FIG. 7. In ovo inhibition assay of N3 against IBV. (A) The effect of N3 (represented by the percentage of uninfected embryos) when it was introduced 3 h after inoculation by a 100-EID$_{50}$ titer of IBV M41 virus. Eight embryos were subjected to each dose of N3. (B) The effect of N3 when it was introduced 6 h after inoculation by a 100-EID$_{50}$ titer of IBV M41 virus. Six embryos were subjected to each dose of N3. (C) The effect of N3 when it was preintroduced 3 h before inoculation by a 100-EID$_{50}$ titer of IBV M41 virus. Eight embryos were subjected to each dose of N3. (D) The effect of N3 when it was preintroduced 6 h before inoculation by a 100-EID$_{50}$ titer of IBV M41 virus. Six embryos were subjected to each dose of N3. The percentage of uninfected is shown in black, and the percentage of infected is shown in gray.

more potent inhibition against SARS-CoV M^{pro} than N3, although they show inhibition against IBV M^{pro} similar to that of N3 (Table 2). Their cocrystal structures with SARS-CoV M^{pro} (data not shown) indicate that the substitution of larger side chains for Val at the P3 position could enhance the van der Waals interaction with the side chain of Glu-166.

DISCUSSION

In this study, we analyzed the structures of native IBV M^{pro} and a SARS-CoV M^{pro} active-site mutant, H41A, in complex with its N-terminal peptide substrate. First, in the structure of IBV M^{pro}, there are three crystallographically independent M^{pro} molecules. Two of them form a symmetric, catalytically active homodimer with two identical but independent active sites. The other M^{pro} molecule is identified with its C-terminal autocleavage residues inserted into one of the substrate binding sites of the catalytic dimer. Second, the substrate-bound structure of the SARS-CoV M^{pro} mutant offers novel, detailed information on the binding pattern of the P6 to P5′ sites. For example, the main chain of the P6 to P1 β strand is at an angle of ~125° to the P1′ to P5′ strand around the P1 to P1′ site, which probably facilitates the specific cleavage at this site.

A comparison of the substrate binding sites in the two M^{pro} structures provides structural insights for the design of substrate-based inhibitors targeting CoV M^{pro}s. The conservation at the S1′ and S2′ subsites in the two structures has prompted us to design a new generation of inhibitors with, for example, a small P1′ residue and a long hydrophilic P2′ side chain. Furthermore, at the substrate binding site of molecule A in IBV M^{pro}, the orientation of the long side chain of Arg-P3 (Arg-C305) is in accordance with that of Glu-A164, which would stabilize the substrate (or inhibitor) binding at this position via a van der Waals interaction between these two residues (Fig. 1C). The interaction discussed above suggests that the modification of the P3 position with a relatively long side chain potentially is beneficial for inhibitor-M^{pro} binding. This notion is strongly supported by the observation that N27 and H16 inhibitors, both of which have a larger side chain than N3 at the P3 position, show significantly improved inhibition ability against SARS-CoV M^{pro}. Moreover, the S2 subsite in IBV

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

FIG. 6. Structure of IBV M^{pro} in complex with N3. (A) Chemical structures of inhibitors N3, N27, and H16. (B) Surface representation of IBV M^{pro} (white) in complex with N3 (magenta). The P1 to P5 and P1′ groups and residues forming the substrate binding pocket are labeled. (C) Stereo view showing N3 bound into the substrate binding pocket of IBV M^{pro}. The N3 inhibitor is shown in gold and is covered by an omit map at 2.0-Å resolution contoured at 1.2 σ. Residues forming the substrate binding pocket are shown in silver. A water molecule (in red) forms hydrogen bonds with N3. (D) Superposition of the substrate-binding pockets of IBV M^{pro}-N3 complex and SARS-CoV M^{pro}-N3 complex (in stereo). The N3 inhibitor bound into the substrate binding pocket of SARS-CoV M^{pro} (cyan) is in yellow, while the N3 inhibitor bound into the substrate binding pocket of IBV M^{pro} (green) is in magenta.

M^{pro} is slightly wider than the corresponding subsite of SARS-CoV, which suggests that the substitution of a larger residue at the P2 position may enhance the efficacy of inhibitors targeting IBV M^{pro}. In addition, according to the native IBV M^{pro} structure, the monomeric form of M^{pro} reveals significant structural flexibility in the interdomain linker region, which may allow incidental activity during CoV M^{pro} maturation. Thus, locking the loop region into a certain conformation may provide a new strategy to block the activity of CoV polyprotein.

In addition to the above-described two structures that show specific substrate/product binding of CoV M^{pro}s, we also characterized the inhibitor-bound structure of IBV M^{pro}, which shares a similar binding mode with the previously reported M^{pro} structures in complex with Michael acceptor inhibitors. The in ovo inhibition assay performed in chicken embryos, together with the in vitro inhibition assay, provides evidence that N3 can block IBV replication via inhibition of the M^{pro}, the key enzyme in the viral replication cycle. N3 shows a high level of inhibitory efficiency, as our results indicate that a 0.64-μmol (or 44-μg) dose of inhibitor per embryo is sufficient to effectively prevent infection by IBV. As reported previously, the Michael acceptor inhibitors are wide-spectrum inhibitors against all CoV-associated diseases. Hence, the in ovo inhibition assay reported here provides a feasible approach for the discovery of anti-SARS drug candidates, which is important considering the high risk to human health posed by SARS-CoV.

ACKNOWLEDGMENTS

We thank Dawei Ma and Wenhua Li for synthesizing the inhibitors; Luhua Lai for supplying the N-terminal 11-peptidyl substrate of SARS-CoV M^{pro}; Sheng Ye, Xiuhua Dong, Xuemei Li, and Zhiyong Lou for technical assistance; and Rongguang Zhang and Andrzej Joachimiak for data collection for Se-Met IBV M^{pro}.

This work was supported by Project 973 of the Ministry of Science and Technology (grant numbers 2006CB806503 and 2007CB914301), Project 863 of the Ministry of Science and Technology (grant number 2006AA02A322), the NSFC (grant numbers 30221003 and 30730022), the Chinese Academy of Sciences Knowledge Innovation Project (grant number KSCX1-YW-R-05), the Sino-German Center [grant number GZ 236(202/9)], the Sino-European Project on SARS Diagnostics and Antivirals (SEPSDA) of the European Commission (grant number 003831), and other programs (grant numbers 5200638, NCET-06-0752, IRT0723, and 2006AA10A205).

REFERENCES

1. **Anand, K., G. J. Palm, J. R. Mesters, S. G. Siddell, J. Ziebuhr, and R. Hilgenfeld.** 2002. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. EMBO J. **21:**3213–3224.
2. **Anand, K., J. Ziebuhr, P. Wadhwani, J. R. Mesters, and R. Hilgenfeld.** 2003. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science **300:**1763–1767.
3. **Brown, T. D., and I. Brierley.** 1995. The coronavirus nonstructural proteins, p. 191–217. *In* S. G. Siddell (ed.), The Coronaviridae. Plenum Press, New York, N.Y.
4. **Brünger, 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. D Biol. Crystallogr. **54:**905–921.
5. **Cavanagh, D.** 2003. Severe acute respiratory syndrome vaccine development: experiences of vaccination against avian infectious bronchitis coronavirus. Avian Pathol. **32:**567–582.
6. **Cunningham, C. H.** 1973. Immunologic methods in avian research: neutralization test. Avian Dis. **17:**227–235.
7. **Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H. D. Klenk, A. D. Osterhaus, H. Schmitz, and H. W. Doerr.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1967–1976.
8. **Emsley, P., and K. Cowtan.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60:**2126–2132.
9. **Fouchier, R. A., T. Kuiken, M. Schutten, G. van Amerongen, G. J. van Doornum, B. G. van den Hoogen, M. Peiris, W. Lim, K. Stohr, and A. D. Osterhaus.** 2003. Aetiology: Koch's postulates fulfilled for SARS virus. Nature **423:**240.
10. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Coronavirus genome: prediction of putative functional domains in the nonstructural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res. **17:**4847–4861.
11. **Hendrickson, W. A., and M. M. Teeter.** 1981. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. Nature **290:**107–113.
12. **Horton, R. M., H. D. Hunt, S. N. Ho, J. K. Pullen, and L. R. Pease.** 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77:**61–68.
13. **Ignjatoviæ, J., and S. Sapats.** 2000. Avian infectious bronchitis virus. Rev. Sci. Tech. **19:**493–508.
14. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, and L. J. Anderson.** 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1953–1966.
15. **Kuiken, T., R. A. Fouchier, M. Schutten, G. F. Rimmelzwaan, G. van Amerongen, D. van Riel, J. D. Laman, T. de Jong, G. van Doornum, W. Lim, A. E. Ling, P. K. Chan, J. S. Tam, M. C. Zambon, R. Gopal, C. Drosten, S. van der Werf, N. Escriou, J. C. Manuguerra, K. Stohr, J. S. Peiris, and A. D. Osterhaus.** 2003. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet **362:**263–270.
16. **Lai, M. M. C., and K. V. Holmes.** 2001. Coronaviridae: the viruses and their replication, p. 1163–1179. *In* D. M. Knipe and P. M. Howley (ed.), Fields virology, 4th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
17. **Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton.** 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26:**283–291.
18. **Lee, H. J., C. K. Shieh, A. E. Gorbalenya, E. V. Koonin, N. La Monica, J. Tuler, A. Bagdzhadzhyan, and M. M. Lai.** 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology **180:**567–582.
19. **Lee, T. W., M. M. Cherney, C. Huitema, J. Liu, K. E. James, J. C. Powers, L. D. Eltis, and M. N. James.** 2005. Crystal structures of the main peptidase from the SARS coronavirus inhibited by a substrate-like aza-peptide epoxide. J. Mol. Biol. **353:**1137–1151.
20. **Li, J., W. Shen, M. Liao, and M. Bartlam.** 2007. Preliminary crystallographic analysis of avian infectious bronchitis virus main protease. Acta Crystallogr. F **63:**24–26.
21. **Murshudov, G. N., A. A. Vagin, and E. J. Dodson.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. **53:**240–255.
22. **Newcomer, M. E., T. A. Jones, J. Aqvist, J. Sundelin, U. Eriksson, L. Rask, and P. A. Peterson.** 1984. The three-dimensional structure of retinol-binding protein. EMBO J. **3:**1451–1454.
23. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode, p. 307–326. *In* C. W. Carter, Jr., and R. M. Sweet (ed.), Macromolecular crystallography, part A, vol. 276. Academic Press, San Diego, CA.
24. **Peiris, J. S., S. T. Lai, L. L. Poon, Y. Guan, L. Y. Yam, W. Lim, J. Nicholls, W. K. Yee, W. W. Yan, M. T. Cheung, V. C. Cheng, K. H. Chan, D. N. Tsang, R. W. Yung, T. K. Ng, and K. Y. Yuen.** 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet **361:**1319–1325.
25. **Pereira, H. G.** 1989. Coronaviridae, p. 42–57. *In* J. S. Porterfield (ed.), ANDREWES' viruses of vertebrates, 5th ed. Bailliere Tindall, London, United Kingdom.
26. **Reed, L., and H. Muench.** 1938. A simple method of estimating fifty percent endpoint. Am. J. Hyg. **27:**493–497.
27. **Schneider, T. R., and G. M. Sheldrick.** 2002. Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. **58:**1772–1779.
28. **Shi, J., Z. Wei, and J. Song.** 2004. Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: defining the extra domain as a new target for design of highly specific protease inhibitors. J. Biol. Chem. **279:**24765–24773.
29. **Snijder, E. J., and W. Spaan.** 1995. The coronaviruslike superfamily, p. 239–255. *In* S. G. Siddell (ed.), The Coronaviridae. Plenum Press, New York, N.Y.
30. **Spaan, W. J. M., and D. Cavanagh.** 2004. Coronaviridae, p. 945–962. *In* C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV

(ed.), Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses. Elsevier-Academic Press, London, United Kingdom.

31. **Terwilliger, T. C.** 2000. Maximum-likelihood density modification. Acta Crystallogr. D Biol. Crystallogr. **56:**965–972.
32. **Terwilliger, T. C., and J. Berendzen.** 1999. Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. **55:**849–861.
33. **Xue, X., H. Yang, W. Shen, Q. Zhao, J. Li, K. Yang, C. Chen, Y. Jin, M. Bartlam, and Z. Rao.** 2007. Production of authentic SARS-CoV M(pro) with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. J. Mol. Biol. **366:**965–975.
34. **Yang, H., W. Xie, X. Xue, K. Yang, J. Ma, W. Liang, Q. Zhao, Z. Zhou, D. Pei, J. Ziebuhr, R. Hilgenfeld, K. Y. Yuen, L. Wong, G. Gao, S. Chen, Z. Chen, D. Ma, M. Bartlam, and Z. Rao.** 2005. Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol. **3:**e324.
35. **Yang, H., M. Yang, Y. Ding, Y. Liu, Z. Lou, Z. Zhou, L. Sun, L. Mo, S. Ye, H. Pang, G. F. Gao, K. Anand, M. Bartlam, R. Hilgenfeld, and Z. Rao.** 2003. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc. Natl. Acad. Sci. USA **100:**13190–13195.
36. **Ziebuhr, J.** 2005. The coronavirus replicase. Curr. Top. Microbiol. Immunol. **287:**57–94.
37. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J. Gen. Virol. **81:**853–879.

Downloaded from http://jvi.asm.org/ on April 25, 2015 by EAST CAROLINA UNIV