9470–9489 *Nucleic Acids Research, 2022, Vol. 50, No. 16* *Published online 10 August 2022*
https://doi.org/10.1093/nar/gkac671

Biochemical and structural characterization of hepatitis A virus 2C reveals an unusual ribonuclease activity on single-stranded RNA

Pu Chen[1,†], Justyna Aleksandra Wojdyla[2,†], Ombretta Colasanti[3,†], Zhijian Li[1,†], Bo Qin[1], Meitian Wang[2], Volker Lohmann[3,*] and Sheng Cui[1,*]

[1]NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, PR China, [2]Swiss Light Source at the Paul Scherrer Institute, CH-5232 Villigen, Switzerland and [3]University of Heidelberg, Department of Infectious Diseases, Molecular Virology Centre for Integrative Infectious Disease Research (CIID) INF 344, 1st. Floor, D-69120 Heidelberg, Germany

Received June 13, 2022; Editorial Decision July 12, 2022; Accepted July 22, 2022

ABSTRACT

The HAV nonstructural protein 2C is essential for virus replication; however, its precise function remains elusive. Although HAV 2C shares 24–27% sequence identity with other 2Cs, key motifs are conserved. Here, we demonstrate that HAV 2C is an ATPase but lacking helicase activity. We identified an ATPase-independent nuclease activity of HAV 2C with a preference for polyuridylic single-stranded RNAs. We determined the crystal structure of an HAV 2C fragment to 2.2 Å resolution, containing an ATPase domain, a region equivalent to enterovirus 2C zinc-finger (ZFER) and a C-terminal amphipathic helix (PBD). The PBD of HAV 2C occupies a hydrophobic pocket (Pocket) in the adjacent 2C, and we show the PBD–Pocket interaction is vital for 2C functions. We identified acidic residues that are essential for the ribonuclease activity and demonstrated mutations at these sites abrogate virus replication. We built a hexameric-ring model of HAV 2C, revealing the ribonuclease-essential residues clustering around the central pore of the ring, whereas the ATPase active sites line up at the gaps between adjacent 2Cs. Finally, we show the ribonuclease activity is shared by other picornavirus 2Cs. Our findings identified a previously unfound activity of picornavirus 2C, providing novel insights into the mechanisms of virus replication.

INTRODUCTION

Hepatitis A virus (HAV) is the sole member of the *Hepatovirus genus* among the *Picornaviridae*, underscoring its uniqueness in this family (1). HAV remains a major pathogen causing communicable liver disease worldwide. Although HAV infection is generally mild, a small portion of HAV patients (particularly those older than 50 years) could die from fulminant hepatitis (2). Available HAV vaccines can provide efficient and lasting immune protection (https://www.who.int/news-room/fact-sheets/detail/hepatitis-a, last accessed date, 22 July 2022) (3). Nevertheless, occasional HAV outbreaks occur even in developed countries with good sanitary and hygienic conditions (4,5). To rapidly contain future outbreaks, besides treating fulminant hepatitis A patients as well as controlling potential vaccine-escape variants, development of effective anti-HAV drugs remains a great value in disease control.

The HAV particle shares features with both standard picornavirus and insect *picorna*-like viruses (6). The icosahedral capsid of HAV (∼27 nm in diameter) is remarkably smooth and lacks canyons, which is consistent with its resistance to high temperature and low pH environments (6). The 7.5 kb genome of HAV is a plus-sense single-stranded RNA comprising a 5′-untranslated region (UTR), a large open reading frame, a 3′-UTR followed by a poly(A) tail (6,7). The open reading frame encodes a large polyprotein (∼250 kDa) that is processed co- and post-translationally by viral and cellular proteases, to yield mature structural proteins (VP1–VP4), nonstructural proteins (2A–2C, 3A–3D) as well as multiple functional precursors (7). Unlike standard picornavirus 2A, the HAV 2A is not a protease; instead it is a part of VP1 (8). The 3C and 3ABC proteins of HAV are proteases responsible for most processing steps (9).

*To whom correspondence should be addressed. Tel: +86 10 67828669; Fax: +86 10 67855012; Email: cui.sheng@ipb.pumc.edu.cn
Correspondence may also be addressed to Volker Lohmann. Tel: +49 6221 56 6449; Fax: +49 6221 56 4570; Email: volker.lohmann@med.uni-heidelberg.de
†The authors wish it to be known that, in their opinion, the first four authors should be regarded as Joint First Authors.

A hallmark of HAV infection is its inability to shut off cellular protein synthesis (10). Whereas many picornaviruses inhibit host mRNA translation by targeting cellular initiation factor 4G (eIF4G), HAV must compete with cellular mRNAs for the translation of its own RNAs (11). To survive this competition, HAV adopts multiple strategies including deoptimized-codon usage, which contributes to its low growth rate and inefficient replication in cell culture (12). The HAV 2C binds the cloverleaf structure of the internal ribosome entry site at 5′ end of the plus-strand RNA, indicating its involvement in translation (13).

Picornavirus 2C is a multifunctional protein essential for nearly all steps in the virus life-cycle (14). It is one of the most conserved proteins among the *Picornaviridae*, which makes the coding region of 2C useful for molecular diagnosis and genotyping (15). Being an atypical picornavirus, HAV harbors a unique 2C distantly related to all other 2C homologs in the *Picornaviridae*. Despite a comparable size comprising 335 amino acids (aa) polypeptide chain, HAV 2C shares poor amino acid sequence identity with the 2C from enterovirus 71 (EV71, 27.9%), enterovirus D68 (EV-D68, 28.2%), poliovirus (PV, 27.3%), foot-and-mouth disease virus (FMDV, 27.2%), coxsackievirus B3 (CVB3, 26.0%), echovirus 30 (Echo30, 26.0%) and human rhinovirus (HRV, 27.5%). Nonetheless, HAV 2C exhibits similar domain organization and key functions during infection like its homologs. HAV 2C is predicted to comprise an N-terminal domain responsible for binding membrane (13), a central ATPase domain essential for virus replication and a C-terminal domain. Similar to other picornaviruses 2Cs, HAV 2C (and its precursor 2BC) can induce intracellular membrane rearrangement to provide a platform for replication-complex assembly (16). One distinguishing feature of HAV replication is that the intracellular membranes required for replication-complex assembly seem to be derived from the outer mitochondrial membranes, not from endoplasmic reticulum membranes as in the case of PV replication complex formation (17,18). Following the N-terminal membrane-binding domain is an ATPase domain belonging to the SF3-helicase family of AAA+ proteins, that is the most conserved domain of HAV 2C. The C-terminal region of HAV 2C lacks the Cys-rich motif that forms a zinc-finger in enterovirus 2C. The function of this domain requires further investigation.

Like other SF3-helicase family members (19), HAV 2C presents ATPase and RNA-binding activities, but duplex unwinding activity has not been reported, in contrast to EV71 2C (20). Both full-length HAV 2C and a truncation mutant lacking the N-terminal membrane-binding motif can bind RNA, implying that the RNA binding site is not in the N-terminal region (13). HAV 2C recognizes the 3′-UTR of negative-strand viral RNA (21), suggesting that it might anchor the negative-strand RNA, the template for positive-strand RNA synthesis, to intracellular membranes.

We previously determined crystal structures of EV71 2C and PV 2C (22,23) lacking the N-terminal membrane binding domain. We demonstrated that enterovirus 2C comprises an ATPase domain followed by a zinc-finger and C-terminal amphipathic helix. In our structures, the C-terminal helix of enterovirus 2C occupies a hydrophobic pocket in the adjacent 2C; this specific 2C–2C interaction is essential for ATPase activity of 2C and virus replication (22,23). We proposed a hexameric-ring model for enterovirus 2Cs, which contains a negatively charged central pore. These findings are consistent with the results originally reported for FMDV 2C, which revealed hexamer formation of FMDV 2C in the presence of RNA (24). Investigating whether those features of enterovirus 2C are conserved in the ancestral HAV 2C is important to understand the molecular evolution of picornaviral 2C proteins.

Here, we characterized the HAV 2C protein structurally, biochemically and functionally. We determined the crystal structure of a soluble HAV 2C fragment containing the entire ATPase domain, a region equivalent to enterovirus 2C zinc-finger (ZFER) and a C-terminal amphipathic helix. The C-terminal helix (pocket-binding domain, PBD) of HAV 2C occupies a hydrophobic pocket (Pocket) of the adjacent 2C in crystal lattice, which mediates the 2C–2C interaction in solution. Performing mutagenesis studies, we demonstrated the PBD–Pocket interaction is essential for 2C self-oligomerization, ATPase activity and HAV replication. In our biochemical assays, we found that HAV 2C exhibits single-strand specific ribonuclease (RNase) activity and prefers poly-U segments; and this RNase activity is independent of 2C′s ATPase activity. We identified several acidic residues essential for the RNase activity of HAV 2C and showed those residues are fundamental to HAV replication. Given these acidic residues reside in a conserved region, we confirmed the 2Cs from EV71, FMDV, CVB3, PV and HRV all share RNase activity. Using a cell based HAV replication assay, we verified the functional importance of ATPase, ZFER, PBD and Pocket of HAV 2C for HAV replication.

MATERIALS AND METHODS

Protein expression and purification

The sequence of the prototype HAV stain HM-175 (GenBank: M14707.1) was used for recombinant 2C expression, biochemical characterization and structure determination. To express a soluble fragment of HAV 2C (128–335 aa, termed HAV 2C-ΔN), the gene of HAV 2C-ΔN was amplified by polymerase chain reaction (PCR) and inserted between NcoI and HindIII restriction sites of a modified pET-32a vector with a tobacco etch virus (TEV) protease cleavage site. The plasmid was transformed to *Escherichia coli* BL21 (DE3) competent cells. The bacteria were cultured in Luria-Bertani (LB) medium at 37°C and induced at 18°C with 0.5 μM isopropyl-beta-D-thiogalactopyranoside (IPTG) when the OD_{600} of the medium reached ∼0.6. The culturing continued for 18 hours after induction. The bacteria cells were harvested by centrifugation 5422g for 10 min. The cell pellet was re-suspended in a lysis buffer containing 20 mM HEPES-Na pH 7.4, 300 mM NaCl and disrupted by ultrasonification. The lysate was clarified by centrifugation at 47 850g at 4°C for 30 min. The supernatant was passed through Ni-NTA resin (GE healthcare) and eluted with the elution buffer containing 20 mM HEPES-Na pH 7.4, 300 mM NaCl and 300 mM imidazole. The eluate was dialyzed overnight at 4°C in the dialysis buffer containing 20 mM HEPES-Na pH 7.4 and 300 mM NaCl. TEV protease was added to the dialysis buffer to remove the 6 × His

tag. Finally, HAV 2C-ΔN was purified with Superdex®200 10/300 GL column (GE healthcare) pre-equilibrated in the gel filtration buffer containing 10 mM HEPES-Na pH 7.4 and 100 mM NaCl.

To express full-length HAV 2C (termed WT HAV 2C), the gene of WT HAV 2C was amplified by PCR and cloned to baculovirus transfer vector pFastbac1. The coding sequence of 6 × His-tag was added to the N-terminal of HAV 2C. WT HAV 2C was overexpressed in sf-21 cells using the Bac-to-Bac Baculovirus Expression System (Invitrogen). A 600 ml cell culture (2.0×10^6 cells ml^{-1}) was infected with 18 ml baculovirus at 28°C. At 48 h post-infection, cells were harvested by centrifugation at 5422g for 10 min. The cell pellet was resuspended in a lysis buffer containing 20 mM HEPES-Na pH 7.4, 300 mM NaCl and lysed by ultrasonification. After centrifugation at 47 850g 4°C for 1 h, the supernatant was discarded and the precipitation was solubilized in the solubilization buffer 20 mM HEPES-Na pH 7.4, 300 mM NaCl and 20 mM *n*-dodecyl β-D-maltoside (DDM, Anatrace, D310) for 3 h at room temperature. The resulting mixture was clarified by centrifugation (47 850g for 2 h) and the supernatant containing WT HAV2C was loaded into the Ni-NTA column (GE healthcare). The protein was eluted with the elution buffer and finally purified using size-exclusion chromatography (Superdex 200 10/300 GL, GE healthcare). WT HAV 2C was concentrated to 10 μM in the buffer containing 10 mM HEPES-Na pH 7.4, 100 mM NaCl and 0.4 mM DDM before stored at –80°C. Plasmids encoding full-length HAV 2C mutants were obtained using site-directed mutagenesis and the mutants were expressed using the similar method described above. SDS-PAGE analysis of all HAV 2C variants is shown in Supplementary Figure S2.

Genes encoding the 2C proteins of enterovirus 71 (EV71), enterovirus D68 (EV-D68), poliovirus (PV), foot-and-mouse disease virus (FMDV), coxsackievirus B3 (CVB3), Echovirus 30 (EcoV30) and human rhinovirus (HRV) were inserted into pMal-c2X plasmid between the BamHI and HindIII sites. These 2C proteins were all fused with an N-terminal maltose-binding protein (MBP) tag. Overexpression of these 2Cs in *E. coli* BL21 (DE3) competent cells followed the same protocols as described above. After the disruption of the bacteria cells, the supernatants of the cell lysate were passed through amylose resin (GE healthcare) and eluted with elution buffer containing 20 mM HEPES-Na pH 7.4, 300 mM NaCl, 10 mM maltose and 2 mM EDTA. All subsequent purification steps followed the protocol for purifying HAV 2C proteins as describe above.

Crystallization and structure determination

The soluble HAV 2C fragment 128–335 aa was concentrated to ∼2.5 mg/ml in a buffer containing 10 mM HEPES-Na pH 7.4 and 100 mM NaCl before crystallization trails. The crystallization of HAV 2C 128–335 aa was achieved by mixing 1 μl of protein with 1 μl of reservoir buffer containing 400 mM $NH_4H_2PO_4$. The crystals were grown in a hanging-drop vapor-diffusion system at 18°C. The crystals were flash frozen in liquid nitrogen before X-ray diffraction experiments.

Table 1. Data collection and refinement statistics

	HAV 2C fragment 128–335aa SAD data (Sulphur) X ray λ = 2.08Å	HAV 2C fragment 123–335aa Native data X ray λ = 1.00Å (PDB: 7XT3)
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	50.03, 54.56, 85.37	50.17, 54.69, 85.67
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution Å	46.02–2.37	43.3–2.15
R_{merge} (last shell)	0.19 (1.73)	0.15 (1.79)
$I / \sigma I$ (last shell)	18.28 (2.82)	7.67 (0.81)
Completeness % (last shell)	100.0 (100.0)	99.9 (99.6)
Redundancy (last shell)	46.93 (45.38)	4.30 (4.38)
Refinement		
Resolution Å	/	2.15
No. reflections (No. reflections in cross validation)	/	24 647 (1221)
R_{work}/R_{free} (last shell)	/	0.22/0.26 (0.36/0.45)
No. atoms		
Protein	/	2692
Ligand/ion	/	15
Water	/	40
***B*-factors**		
Protein	/	59.71
Ligand/ion	/	67.24
Water	/	50.35
R.m.s. deviations		
Bond lengths (Å)	/	0.005
Bond angles (°)	/	0.714

Given HAV 2C shares limited amino acid sequence similarity with other picornaviral 2C proteins, suitable model for molecular replacement was unavailable. The Native-SAD method (25) was employed for *ab initio* phasing. The soluble fragment of HAV 2C contains 7 cysteine and 10 methionine in 208 residues, which provides total 17 sulphur atoms to contribute to anomalous signal. Seven complete datasets were collected for a single crystal using the X-ray with the wavelength of 2.08 Å. The crystal diffracted the X-ray to 2.37 Å, belonged to the space group of $P2_12_12_1$ and contained a single copy of HAV 2C in asymmetric unit. The chi angle of these datasets ranged from 0° to 30°, and the redundancy reached ∼38. After scaling and merging of these datasets, the anomalous signal extended to ∼3.0 Å. An initial electron density map was then calculated using the software SHELXC/D/E (26), which allowed manual building of an initial model using software Coot (27). To improve resolution, we collected reflections of another crystal using higher X-ray energy, wavelength 1.0 Å. This crystal diffracted the X-ray to 2.15 Å. The crystal structure was solved by molecular replacement using software Phaser MR (28) with the initial model from Native-SAD as the searching model. The high-resolution structure was finally refined using software Phenix (29) (Supplementary Figure S3). The data collection, refinement parameter statistics of the crystal structures of HAV 2C soluble fragment are summarized in Table 1.

Nuclease assay

All RNA and DNA strands used for nuclease assay and helicase assay were synthesized in Sangon Biotech (Shanghai). The sequence of the RNA strands is listed in below table:

Oligo name	Sequence	Modification	Length (nt)
RNA1	Cy3-5′-CGAAGCUGCUAACAUCAG-3′	Cy3	18
RNA2	5′-CUGAUGUUAGCAGCUUCG-3′	None	18
RNA3	5′-CUGAUGUUAGCAGCUUCG UUUUUUUUUUUUUUU-3′	None	33
RNA4	5′-UUUUUUUUUUCUGAUGUU AGCAGCUUCG-3′	None	28
RNA5	Cy3-5′-GGAGUGGUGCUUUUUG UCAUAAGGUUGGUGUCAGG-3′	Cy3	35
RNA6	Cy3-5′-GG AGUGGUGCUUUUUGUCAUAA GGUUGGUGUCAGG-3′	Cy3, 2′-O-Me at uridines	35
RNA7	Cy3-5′-CUGAUGUUAGCAGCUU UUUUUUUUUUUUUUUUUU UUUUUUUU-3	Cy3	42
RNA8	Cy3-5′-CUGAUGUUAGCAGCAA AAAAAAAAAAAAAAAAAA AAAAAA-3′	Cy3	42
RNA9	Cy3-5′-CCUGACACCAACCUUA UGACUUUUUGCACCACUCC-3′	Cy3	35
DNA1	Cy3-5′-GGAGTGGTGCTTTTTG TCATAAGGTTGGTGTCAGG-3′	Cy3	35

RNA5, RNA6 and RNA9 are both 35 nt single strand RNAs; RNA5 and RNA9 were mixed to anneal for dsRNA35 with five U mismatched in the middle of the sequence; RNA1 and RNA2 was annealed to form the dsRNA18 matched completely; RNA1 annealed with RNA3 with 3′ terminal overhang while with RNA4 with 5′ overhang.

Nuclease assay: reactions contained 1 μM HAV 2C-ΔN (or 1 μM full-length HAV 2C, or one of its mutant derivatives) and 0.1 μM RNA (or DNA). Reactions were performed in 25 mM HEPES-Na pH 7.4, 80 mM NaCl, 35 mM KCl, 5 mM $MgCl_2$, 1 mM DTT and RNaseOUT™ Ribonuclease Inhibitor (40 unit per 20 μl reaction mixture). The Ribonuclease Inhibitor was added to reaction mixtures 5 min before the addition of 2C protein. Eppendorf tubes, filter-tips and other consumables were all RNase-free grade. After incubation at 37°C for a certain time, the reactions were stopped by addition of an equal volume of loading buffer (50% glycerol, 1 × TBE, 10 mM EDTA). The products were then analyzed in urea-containing 15% polyacrylamide gels (www.wshtbio.com, E301U15F) or 15% native-polyacrylamide gels.

Helicase assay

Helicase assay was carried out as described previously (30). Briefly, the 10 μl reaction mixtures contained 50 mM HEPES-Na pH 7.4, 5 mM $MgCl_2$, 2 mM DTT, 1 mM ATP, 50 nM of partial duplex RNA substrate (RNA4 & RNA1) and 300 nM unlabelled trap RNA (RNA2). The reaction was initiated by adding 1 μM enzymes (MERS-CoV nsp13, HAV 2C FL or HAV 2C ΔN) and incubated at 37°C for 30 min. Reactions were terminated by adding 2.5 μl loading buffer containing 100 mM Tris–HCl pH 7.5, 1% SDS, 50 mM EDTA and 50% glycerol. All reaction mixtures were then resolved by Native-PAGE. The gels were visualized using Typhoon Trio Variable Mode Imager (GE healthcare).

In vitro transcription, electroporation of HAV subgenomic RNAs and detection of luciferase activity

For production of *in vitro* transcripts 10 μg of the respective plasmid were linearized for 60 min at 37°C with restriction enzyme AgeI. Production of *in vitro* transcripts and electroporation were described previously (31). For electroporation, 1×10^7 Huh7-Lunet cells were resuspended in 1 ml cytomix. Cytomix refers to a buffer that resembles the intracellular ionic composition, which has been previously shown to improve the efficiency of transfection and cell viability during electroporation (32). 10 μg of subgenomic replicon RNA were used for transfection of 400 μl of single cell suspension. Cells were seeded and harvested as described before (31). The protocols for measurement of Firefly and Renilla-luciferase activity were described elsewhere (31). Note that luciferase activity at early time points after transfection (here: 4 h) reflects translation of the input RNA, which can be influenced by variations in transfection efficiency or by RNA stability. Since we did not observe consistent differences among mutants at this time pointing to systematic effects, we used these data to normalize for variations in transfection efficiency.

EV71 virus titer measurement

EV71 infectious clone (pEV71) contained the full-length cDNA of WT virus, and mutations in the 2C region (K135A and D186A) were introduced by the site-directed mutagenesis. To produce RNA transcripts, 8 μg of pEV71 plasmid or mutants DNAs were linearized by Hind III digestion for 60 min at 37°C, and the DNAs were *in vitro* transcribed into RNAs using the MEGA script T7 Kit (Ambion). Equal amounts of RNAs were then transfected into Vero cells using Lipo2000 transfection reagent (Invitrogen). Cells were collected at 72 h post-transfection, and were freeze-thaw lysed. The cell lysates were centrifuged at 12 000g for 5 min at 4°C, and the supernatant containing the viruses were collected for titer measurement. 10-fold serial dilution of the supernatant were carried out to prepare virus solutions, and 100 μl of virus solutions was used for infecting RD cells seeded in 96-well plate. The virus titers were measured by $TCID_{50}$ method in RD cells (Behrens–Kraber algorithm) by counting the wells in which CPE appeared, using equation: $TCID_{50} = 10^{1+d(s+0.5)}$, where d indicates $\log_{10}$ of the dilution factor, s indicates the sum of the ratio of cytopathic effect (CPE). In every single assay, 1×10^4 RD cells were seeded in 96-well plate, 24 h after seeding, ten groups of diluted viruses were used to infect RD cells, eight repeats were set in every single group. All three assays are triplicated independently.

Plasmid constructs and mutagenesis.

Replication-competent HAV firefly-luciferase reporter replicon pT7-18f-LUC and replication-deficient pT7-18f-LUCmut plasmids (3D mut) (33), derived from the

cell culture adapted HAV strain HM-175 18f (GenBank: M59808), were used to produce subgenomic reporter replicon RNA referred to as HAV wt or HAV mut. PCR-directed mutagenesis was used for generation of 2C mutants, with XhoI and AlwNI restriction sites flanking the 2C gene, basing on the following list of primers:

Construct	Primer sense (5′- 3′)	Primer antisense (5′- 3′)
E169A	TGTTGAGCCTGCGA AGAATAT	ATATTCTTCGCAGGCT CAACA
H276A	TTTTCAAAAATCCTGC AAATGATAT	ATATCATTTGCAGGAT TTTTGAAAA
S206A	CAGATGAGGACTGG GCGGATTTT	AAAATCCGCCCAGTCC TCATCTG
K267A	GGTTGAAGTTGCACCT GCTTCATTT	AAATGAAGCAGGTG CAACTTCAACC
K149A	GAGGGGGAGGAGCG AGCTTAACATC	GATGTTAAGCTCGCTC CTCCCCCTC
L151A	GAGGAAAGAGCGCA ACATCAATTGC	GCAATTGATGTTGC GCTCTTTCCTC
A155R	CTTAACATCAATTCGA TTGGCAACCAAAATTT G	CAAATTTTGGTTGCCA ATCGAATTGATGTTAA G
D194A	GTTTGCATCATTGCTG ATATTGGCCAAAAC	GTTTTGGCCAATATCA GCAATGATGCAAAC
N241A	GCAACTTCAGCTTGGT CAAATCCAAG	CTTGGATTTGACCAAG CTGAAGTTGC
R259A	GCAATTGACCGCGC ACTCCATTTCAAG	CTTGAAATGGAGTG CGCGGTCAATTGC
F271A	GTTAAACCTGCTTCAG CTTTCAAAAATCCTC	GAGGATTTTTGAAA GCTGAAGCAGGTTT AAC
F272A	GTTAAACCTGCTTCAT TTGCCAAAAATCCTC	GAGGATTTTTGGCA AATGAAGCAGGTTT AAC
L284A	GTTGAATGTTAATGCA GCTAAAACAAATG	CATTTGTTTTAGCTGC ATTAACATTCAAC
I291A	CAAATGATGCAGCC AAAGATATGTC	GACATATCTTTGGCTG CATCATTTG
F329A	AACTCAGCTACTGCAC TATCATTATCATCTGA AATTCCCTGAGACCAC AACTCCATGGCTTCAG TCATG	AACTCAGCTACTGCAC TATCATTATCATCTGA ATTCCCTGAGACCAGA ATTCCATGGCTTCAGT CATG
W333A	AACTCAGCTACTGCAC TATCATTATCATCTGA AATTCCCTGAGACGCC AACTCCATG	AACTCAGCTACTGCAC CTATCATTATCATCTG AAATTCCCTGAGACCA CAACTCCATG
D202A	AACACAACAGCAGA GGACTGGTCAG	CTGACCAGTCCTCTGC TGTTGTGTT
E203A	AAACACAACAGATG CCGACTGGTCA	TGACCAGTCGGCATCT GTTGTGTTT
D204A	ACAACAGATGAGGC CTGGTCAGATT	AATCTGACCAGGCCTC ATCTGTTGT
D207A	TGAGGACTGGTCAG CATTTTGTCAG	CTGACAAAATGCTGAC CAGTCCTCA
D180A	CTGTGGCTTCAGCTTA CTGGGATGG	CCATCCCAGTAAGCTG AAGCCACAG
D183A	AGATTACTGGGCTG GATATAGTGGA	TCCACTATATCCAGCC CAGTAATCT

ATPase assay

ATPase assays were performed as previously described (22,23,34,35). The concentration of the enzyme was kept constant to 5 μM in all reactions. The volume of the of the reaction mixtures was 40 μl, contained 20 mM HEPES-K pH 7.4, 4 mM magnesium acetate, 2 mM dithiothreitol (DTT), 500 μM ATP mixed with trace amount of radioactive labelled ATP (γ-^{32}P). The mixtures were incubated at 30°C and the reactions were initiated by adding enzyme. After incubation for 10 min, 20 μl of the reaction mixture was added into the Eppendorf tube with pre-added EDTA (the final concentration is 0.1 M) on ice to quench the reaction. Finally, 1 μl of the sample was spotted on the thin-layer chromatography cellulose TLC plates (Sigma-Aldrich) and resolved with running buffer containing 0.8 M acetate and 0.8 M LiCl. The plates were dried and analysed using storage phosphor screen and Typhoon Trio Variable Mode Imager (GE healthcare). ATP turnover was quantified using ImageQuant TL software (GE Healthcare). The buffer conditions used here for HAV 2C ATPase assay are the same as the one used for PV 2C ATPase assay, previously reported (22), which allows us for activity comparison among different 2C proteins. Of note, mild detergent (DDM) was included in HAV 2C FL protein for stabilization, whereas no detergents were present in PV 2C protein preparation.

Cell culture

Huh7-Lunet cells (31), a subclone of the human hepatoma cell line Huh7, were grown in Dulbecco's modified minimal essential medium (DMEM; Life Technologies, Norwalk, CT), supplemented with 2 mM of L-glutamine, nonessential amino acids, 100 U/ml of penicillin, 100 U/ml of streptomycin, and 10% fetal calf serum. Vero cells were grown in minimum essential medium (MEM, HyClone) supplemented with 10% fetal calf serum. RD cells were grown in DMEM medium (HyClone) supplemented with 10% fetal calf serum. Sf-21 cells, which is derived from the pupal ovarian tissue of the fall army worm, *Spodoptera frugiperda*, was grown in Sf-900™ II SFM medium (Invitrogen), supplemented with 50 mg/l gentamycin.

Biolayer interferometry

The experiment was conducted at 25°C in HBSP buffer (10 mM HEPES-Na pH 7.5, 100 mM NaCl, 0.005% surfactant P20) by an Octet RED96e instrument (Fortebio). The volume for the solutions were 200 μl, and the assay was performed in black, solid 96-well, flat bottom plates with a shaking speed of 1000 rpm. The streptavidin (SA) biosensors were activated in double distilled water for at least 10 min. A time of 90 s was performed for sensor check before loading the biotin labeled ssRNA, which took 120 s. After loading process, the biosensors went to the next step of baseline for 120 s until the baseline become flat. A time of 50 s was taken for protein association in the following step. The dissociation was the last step and performed for 50 s. The reference baseline was recorded for a sensor loaded with ligand but no analyte which used to be subtracted to correct any systematic baseline drift. The ssRNA with biotin labeled was synthesized from Songon Biotech (Shanghai) Co., Ltd, denoted as biotin-ssRNA$_{35}$: GGAGUGGUGC UUUUUGUCAUAAGGUUGGUGUCAGG. 0.2 μM of biotinylated ssRNA were immobilized onto the surface of SA biosensors. To get a precise Kd of the interaction, a gradient concentration of HAV 2C-ΔN-D207A (2.18, 4.36, 8.72, 17.5 and 35 μM) was used to associate with the peptide coated SA biosensors. The control group biosensors coated with ssRNA$_{35}$ was exposed into HSBP buffer without HAV

2C-ΔN-D207A. All data were processed by Data Analysis 11.1.

RESULTS

HAV 2C exhibits weak ATPase activity

To prepare protein for structural and biochemical characterization, we ectopically expressed the 2C protein encoded by the prototype HAV strain HM-175 (36). Overexpressing wild-type HAV 2C (denoted WT HAV 2C) in *Escherichia coli* yielded insoluble proteins (Figure 1A, left), probably attributed to the presence of N-terminal membrane binding domain and/or protein misfolding in bacteria. Next, we expressed WT HAV 2C in sf-21 insect cells. We increased the solubility of the protein by supplementing the mild detergent *n*-dodecyl β-D-maltoside (DDM) during purification (Figure 1A, middle). This insect-cell-derived protein did not crystallize. Next, we sought to identify HAV 2C fragments suitable for structural investigation. According to our previous experience, crystallizable 2C fragments (such as EV71 2C and PV 2C etc.) typically lack the N-terminal membrane binding motif but retain intact ATPase domain. Therefore, we carried out progressive deletion of N-terminal portion of HAV 2C up to the Walker A motif, yielding a series of truncations. While many HAV 2C fragments with N-terminal deletion were insoluble upon overexpression in *E. coli*, the fragment encompassing residues 128–335 of HAV 2C was highly soluble. Thus, we used this HAV 2C fragment lacking the first 127 amino acids (aa), denoted HAV 2C-ΔN, for structural characterization (Figure 1A, right), and it eventually yielded protein crystals. We assessed the ATPase activity of both proteins. WT HAV 2C exhibited a turnover number $k_{cat} = 0.0048\ s^{-1}$ in ATP hydrolysis (Figure 1B, left), which was about 100-fold lower than PV 2C, $k_{cat} = 0.46\ s^{-1}$ (23), indicating an unusually slow hydrolysis rate. The ATPase activity of HAV 2C-ΔN was similar to those of the ATPase-inactive mutants K149A (Walker A), D194A (Walker B) and N241A (Motif C, Figure 1B, right), indicating HAV 2C-ΔN lacks ATPase activity.

Crystallization and structure determination of HAV 2C-ΔN

To reveal the structural basis underlying the function of HAV 2C, we carried out crystallographic investigations. Because WT HAV 2C failed to crystallize, we focused on HAV 2C-ΔN, which encompasses the entire ATPase domain and the C-terminal domain, but lacks the N-terminal membrane binding motif. We grew HAV 2C-ΔN crystals in a buffer containing 0.4 M $NH_4H_2PO_4$ in a vapor diffusion system at 18°C. We collected a native dataset but could not solve the structure by molecular replacement using 2C structures available in public databases. To solve the 'phasing' problem, we expressed a derivative containing selenomethionine, but it still failed to crystallize. Then, we employed a native single-wavelength anomalous dispersion method (25), because HAV 2C-ΔN harbors 10 methionines and 7 cysteines in total 208 aa, which provides 1 sulphur per ∼12 residues. We collected seven datasets (360° each) on one crystal at a single position and at different orientations (chi angles ranging from 0° to 30°) using X-ray beams with a wavelength of 2.075 Å. We solved the structure using programs SHELX C/D/E, which found 17 sites with occupancy >0.3, CFOM = 52.62 and yielded an interpretable initial electron density map. We manually built a preliminary model of HAV 2C-ΔN, and we use this model as the molecular replacement searching model to solve another native dataset of higher quality than our initial dataset. We refined our final model, which contains a single HAV 2C-ΔN molecule in the asymmetric unit, to 2.2 Å resolution. Statistics of data collection, integration, scaling and structure refinement are summarized in Table 1.

Overall structure of HAV 2C-ΔN

HAV 2C-ΔN exhibits a domain organization like those of enterovirus 2Cs despite low amino acid sequence identity. It comprises a canonical ATPase domain, a loop topologically equivalent to the zinc-finger of enterovirus 2C (ZFER) and a C-terminal helical domain (Figure 1C). The ATPase domain comprises a five-stranded parallel β-plane (β1–β5) sandwiched by short helices (α1 and α2) and the connecting loops on both sides of the β-plane. Loops connecting β3–α2, α2–β4 and β4–β5 are partially disordered in this structure, indicating their intrinsic flexibility. The Walker A motif $_{143}$-*G*KRGG*GKS*-$_{150}$ constitutes the phosphate-binding loop (P-loop) between β1–α1. Two phosphate groups from the solvent occupy the P-loop. The Walker B motif $_{190}$-VCII*DD*-$_{195}$ lies at the end of the β3-sheet and the downstream loop. Motif C, a SF3-specific motif, lies in the loop downstream from the β4-sheet. Additionally, HAV 2C has another SF3-specific motif known as B' motif. The B' was originally identified at the central pore of the hexameric ring structures of SF3 DNA helicases (37). B' motif of SF3 DNA helicases usually contains multiple positively charged residues that interact with DNA substrates. By contrast, the B' of HAV 2C harbors a conserved segment $_{225}$-EEK-$_{228}$ that is rich in negatively charged residues. The B' motif of HAV 2C is in the α2-to-β4 loop, which is disordered in our structure. Of note, the 'EEK' segment is highly conserved in Picornavirus 2Cs (Supplementary Figure S1).

Our structure reveals that the ZFER of HAV 2C does not form a zinc-finger like those found in enteroviral 2Cs, presenting a major variation among different 2Cs. HAV 2C ZFER spans only 30 residues, whereas the topologically equivalent zinc-finger in PV 2C and EV71 2C contains 41 residues each (Figure 1D). Structural superimposition of HAV 2C ZFER with enterovirus 2C zinc-fingers show that the region is substantially shorter in HAV 2C and lacks the essential residues for zinc coordination. Since HAV 2C ZFER is considerably smaller and more compact, it probably doesn't require zinc coordination for structural stability.

Unlike the single, long helix at the C-terminus of EV71 2C and PV 2C, the C-terminal helix (α6) of HAV 2C is broken into two helices, α6A and α6B, which are oriented nearly perpendicularly to each other, forming a 'knuckle' (Figure 1E). Although this severely bent α6-helix of HAV 2C was probably caused by crystal packing, it indicates the intrinsic flexibility of this helix. This is reminiscent of EV71 2C crystal structure, which revealed that its C-terminal helix undergo rotation around an internal hinge (22).

Using the Dali server, we identified structural homologs of HAV 2C from the Protein Data Bank. Our top hits in-

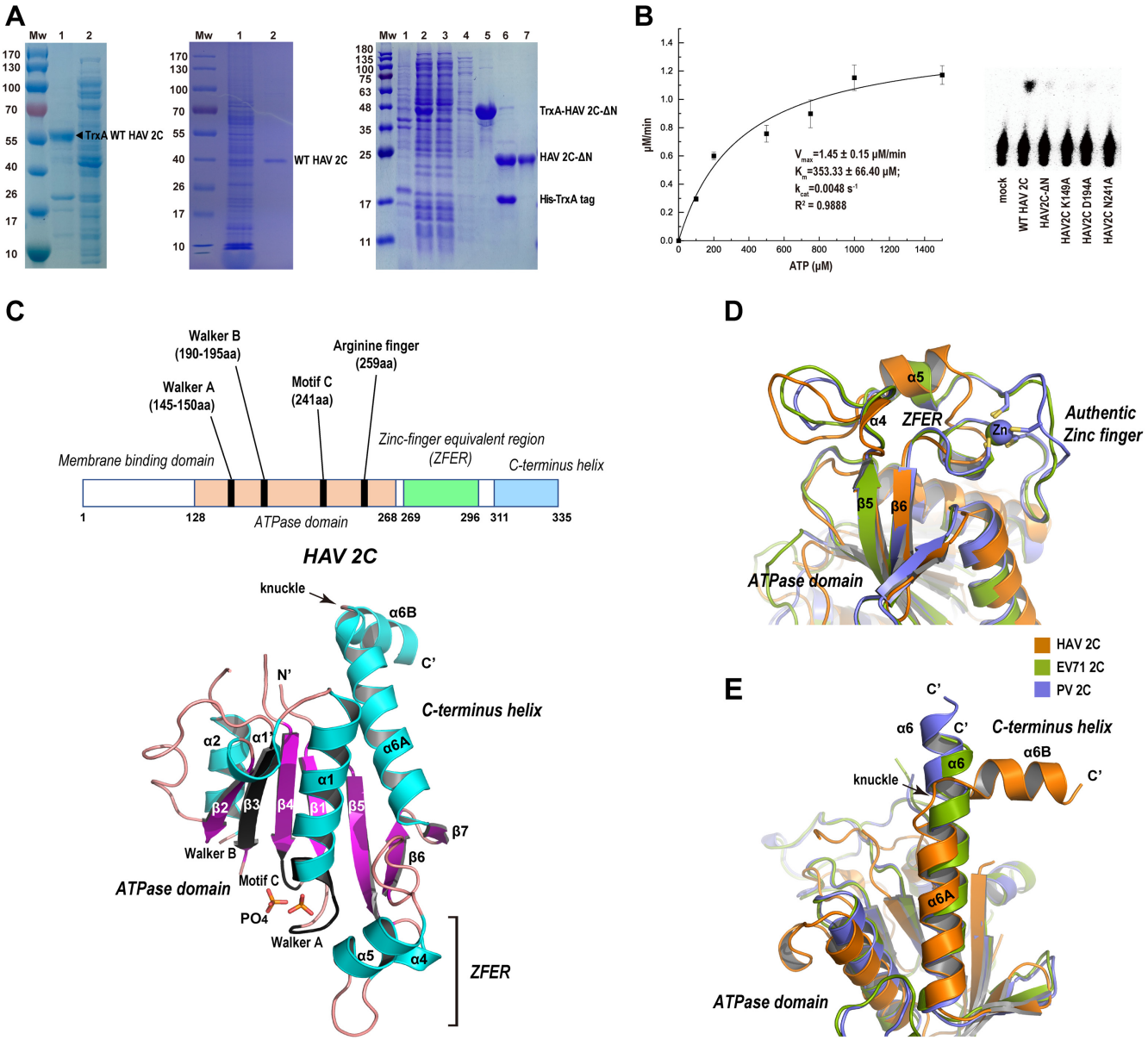

Figure 1. Overall structural feature of HAV 2C-ΔN. (**A**) Expression and purification of WT HAV 2C and HAV 2C-ΔN. Left, expression of WT HAV 2C in bacteria yielded insoluble proteins. Mw: molecular weight standards. The molecular weight of the standard is indicated on the left in kDa. Lane 1, insoluble fraction of the bacteria overexpressing WT HAV 2C; lane 2, supernatant of the bacteria lysate. Middle, production of soluble WT HAV 2C in the presence of DDM. Lane 1, supernatant of sf-21 cell lysate overexpressing WT HAV 2C; lane 2, elution from Ni-NTA resin. Right, production of HAV 2C-ΔN. Mw: molecular weight standards; lane 1, pellet of bacteria lysate; lane 2, supernatant of bacteria lysate; lane 3, flow through of Ni-NTA resin; lane 4, wash of unbound materials; lane 5, sample eluted from Ni-NTA resin; lane 6, Ni-NTA elution digested by TEV protease; lane 7, flow through of the sample from lane 6 reloaded on Ni-NTA. (**B**) Left, the rate of ATP hydrolysis by WT HAV 2C is plotted as the function of ATP concentration; the data was fitted with nonlinear regression using Michaelis–Menten equation to calculate kinetic parameters; K_m, V_{max}, k_{cat} and R^2 of the fitting are indicated. ATPase activity of WT HAV 2C, three ATPase-dead mutants harboring mutation K149A, D194A and N241A and HAV 2C-ΔN. (**C**) Top, diagram of domain/motif organization of HAV 2C protein; key features are indicated. Bottom, ribbon model of HAV 2C-ΔN structure colored by secondary structural elements with annotation; α-helix colored cyan and β-sheet colored magenta. (**D**) Structural superimposition of HAV 2C ZFER (orange ribbon) with the zinc-fingers of PV 2C (blue ribbon) and EV71 2C (green ribbon). The zinc-coordinating residues of two enterovirus 2C proteins are shown with the stick model, demonstrating they are completely shortcut in HAV 2C ZFER. (**E**) Structural superimposition of the C-terminus (α6) of HAV 2C, PV 2C and EV71 2C. The C-terminus helix of HAV 2C breaks down in to two helices (α6A and α6B) that are nearly perpendicular to each other.

cluded enterovirus 2C proteins (38). Our structural superimposition of PV 2C and HAV 2C gave a *Z*-score of 16.7 and aligned 204 Cα atoms with a root-mean-square deviation (RMSD) of 3.6 Å. Structural superimposition of EV71 2C and HAV 2C gave a *Z*-score of 16.5 and aligned 185 Cα atoms with an RMSD of 3.7 Å. This explains the failure in our molecular replacement using either PV 2C or EV71 as a searching model.

The C-terminal helical domain mediates self-oligomerization of HAV 2C

Like other picornavirus 2Cs, HAV 2C self-oligomerized both in crystals and in solution (Figure 2). Crystal packing analysis revealed that all HAV 2C molecules in the crystal were connected to one another via a specific interaction (Figure 2A), forming continuous polymers extending across the lattice. The interface between HAV 2C-2C differs from enterovirus 2C–2C interaction and the HAV 2C interface involves a larger molecular surface, which involves two contact sites: site 1 involves the helix α6A of one 2C and the α1'-β2 region of an adjacent 2C; site 2 involves the C-terminal α6B helix of one 2C and a hydrophobic pocket formed between the ATPase domain and ZFER of an adjacent 2C (Figure 2B and C).

To distinguish which of these sites mediates HAV 2C–2C interaction in solution, we prepared 2C mutants harboring different mutations at these sites. We examined self-oligomerization of these mutants by size-exclusion chromatography (Figure 2D). HAV 2C-ΔN eluted from a Superdex 200 10/300 GL column with an apparent molecular mass ∼160 kDa, indicating that HAV 2C-ΔN formed high-ordered oligomers. HAV 2C-ΔN lacking the C-terminal α6B helix exhibited an apparent molecular mass ∼23 kDa, matching the expected size of a monomer, 23.5 kDa. This result indicates the α6B helix (contact site 2) is essential for 2C self-oligomerization. Further, we found that alanine substitution of residue W333, L332, M330 and F329 on the hydrophobic side of α6B all severely disrupted self-oligomerization. These mutants broken down into dimers or monomers in solution (Figure 2D). By contrast, M317A and Y173A at contact site 1 did not significantly affect 2C self-oligomerization. Both mutants eluted as high-oligomeric species with apparent molecular masses ∼146 kDa and 140 kDa, which were only slightly lower than the high-order HAV 2C-ΔN oligomers (Figure 2D). Collectively, these results indicate that the contact site 2 between α6B helix of one 2C molecule and a hydrophobic pocket on an adjacent 2C is critical for HAV 2C oligomerization in solution while contact site 1 was a result of crystal packing. For simplicity, we denoted the hydrophobic face of the amphipathic α6B helix as the Pocket-Binding Domain (PBD), and the hydrophobic pocket formed between the ATPase domain and ZFER as the Pocket.

Comparing HAV 2C–2C interaction with that of enterovirus 2C, we found that the mode of PBD–Pocket interaction remains similar (Figure 3), although the orientations between 2C–2C are different. This PBD–Pocket mediated 2C oligomerization may present another conserved feature among the *Picornaviridae* 2Cs; thus, agents targeting this interaction can be used as a potential pan-picornavirus antiviral strategy. Although crystal packing was partly responsible for different 2C-2C orientations, intrinsic flexibility of the C-terminal helix is probably common in enterovirus 2Cs and HAV 2C. It is worth noting that the PBD of enterovirus 2Cs usually contains a critical aromatic residue, the penultimate phenylalanine, essential for 2C functions (22,39). By contrast, The PBD of HAV 2C contains two aromatic residues F329 and W333, implying potential druggability at this site.

Impact of 2C mutations on ATPase activity and viral replication

To validate our structural characterization of HAV 2C, we first measured ATP hydrolysis rate of a series of mutants (Figure 4A). Our *in vitro* ATPase assay was hampered by low signal-to-noise ratio despite our efforts to optimize the reaction conditions, including the addition of RNA (Supplementary Figure S4A). Still, the dynamic range of our assays was enough to determine mutations that abrogate the activity. We introduced mutations to full-length HAV 2C by site-directed mutagenesis and measured ATP hydrolysis rate at a constant enzyme concentration (5 μM). We selected E169A, S206A, K267A and H276A as irrelevant mutations because these mutations are not in motifs critical for HAV 2C activities; i.e. we avoided selecting residues within ATPase motifs, ZFER, Pocket and PBD. HAV 2C harboring irrelevant mutations all exhibited a hydrolysis rate similar to wild-type level. On the other hand, we introduced K149A (Walker A), D194A, D195A (Walker B), N241A (Motif C) and R259A (R-finger) to conserved ATPase motifs. As anticipated, these ATP-inactive mutants exhibited 4–12 folds reduction in hydrolysis. Mutations at PBD or Pocket of 2C all led to ATPase activity abrogation because they exhibited a hydrolysis rate similar to that of the ATPase-inactive mutants. This is consistent with our structural analyses (see below sections), which demonstrated that the C-terminus mediated HAV 2C-2C interaction is essential for forming intact ATPase active site. By contrast, HAV 2C mutants K292A, M294A and L284R (ZFER) retained substantial ATPase activity.

To determine the impact of ATPase domain, ZFER and PBD on HAV replication, we employed a cell-based replication assay (Figure 4B–E, Supplementary Figure S4). We used a subgenomic HAV replicon (33) encoding a firefly luciferase reporter for assessing replication efficiency by quantifying luciferase activity. We constructed a selection of HAV replicon plasmids harboring various mutations in the coding sequence of 2C; the plasmids were transcribed and transfected to Huh7-Lunet cells by electroporation. Replicons containing irrelevant mutations did not show impaired replication efficiency, or even increased it in case of K267A, confirming their lack of contribution to vital functions of 2C. In contrast, all mutations within the conserved ATPase motifs resulted in a substantial reduction or abrogation of viral replication 72 h after transfection. In addition, even those mutations in ZFER, PBD and Pocket with moderate or no effects on the ATPase activity of HAV 2C, like L284A, all resulted in a complete loss of replication capability of the replicon (Figure 4C–E, Supplementary Figure S4B–D), arguing for the essential function of these structural fea-

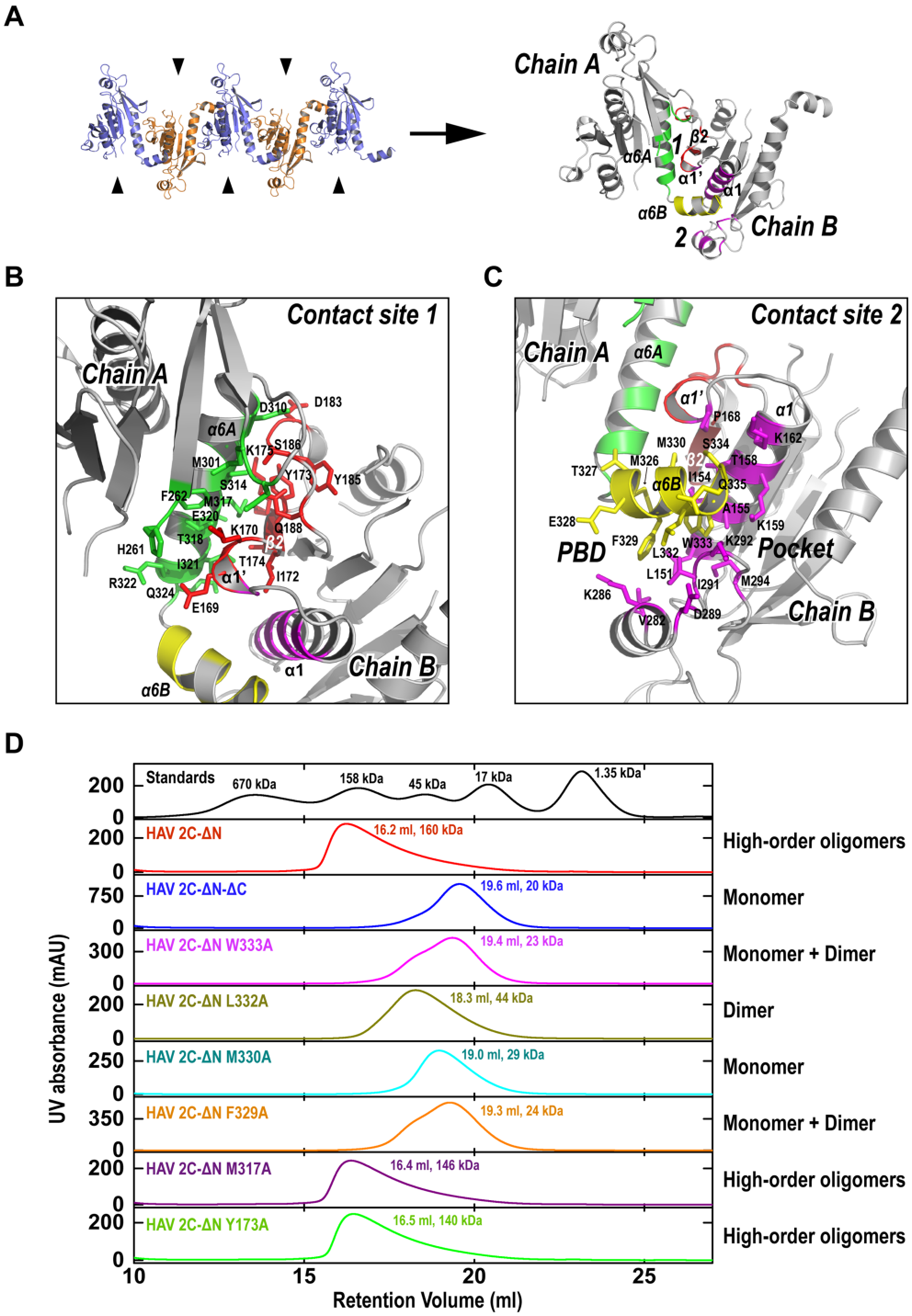

Figure 2. HAV 2C undergoes C-terminus helix dependent self-oligomerization. (**A**) Crystal packing analysis of HAV 2C-ΔN crystals reveals all 2C copies in the lattice are connected via a specific interaction involving two contact sits, indicated as '1' and '2′. Site 1 is between α6A and α1'–β2 loop of the adject 2C, and site 2 is between α6B (termed PBD) and a hydrophobic pocket (termed Pocket) on the adject 2C molecule. (**B**) Magnified view of contact site 1; residues implicated in the interaction are labelled and shown with a stick model. (**C**) Magnified view of contact site 2; residues implicated in the interaction are labelled and shown with a stick model. (**D**) Size-exclusion chromatography profiles of HAV 2C-ΔN and a selection of mutants/truncation harboring mutations at the contact sites. Retention volume from Superdex 200 10/300 GL column, calculated molecular mass and oligomeric state of the proteins are indicated.

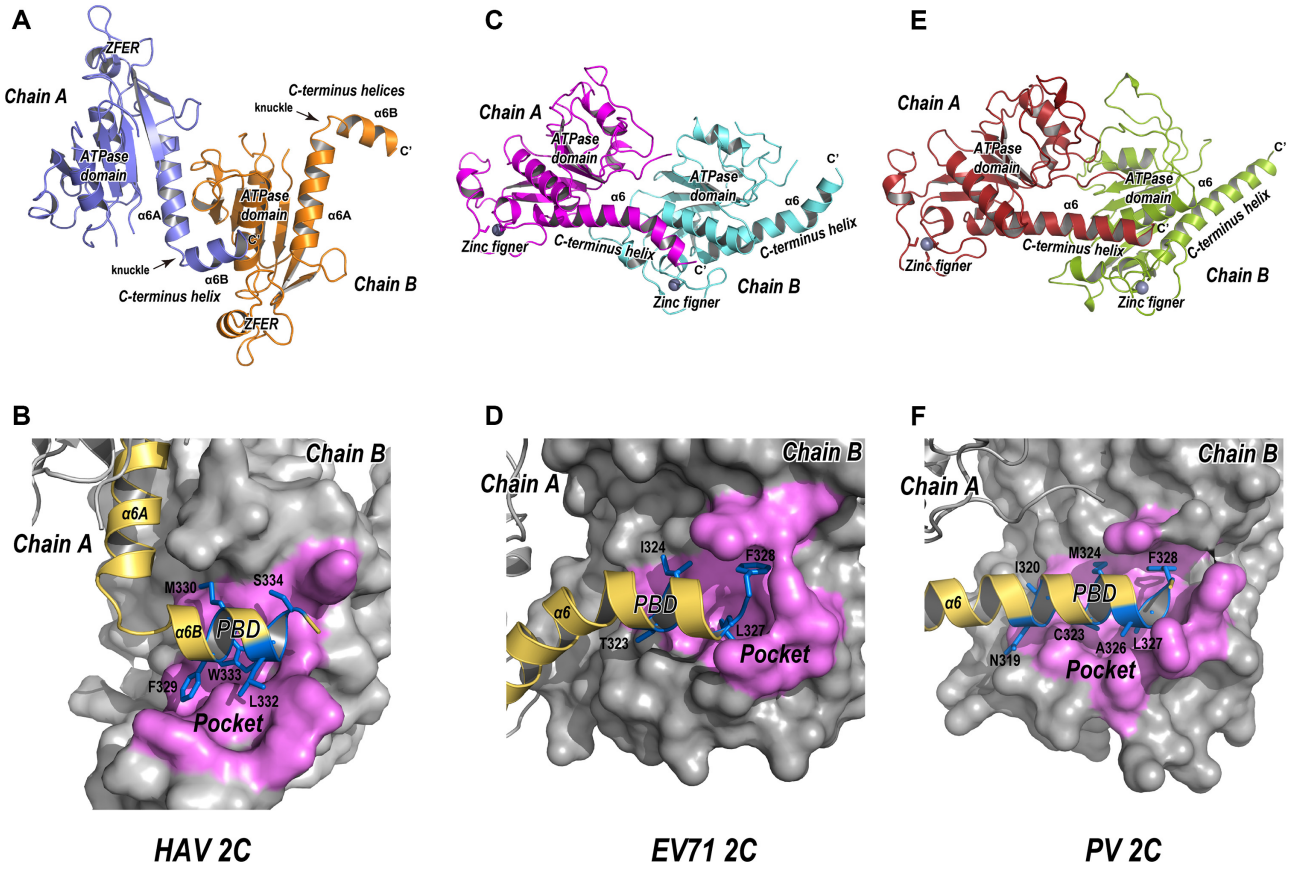

Figure 3. Comparison of the PBD–Pocket interaction in the crystal structures of HAV 2C, EV71 2C and PV 2C. (**A**, **C** and **E**) Ribbon model of HAV 2C dimer, EV71 2C dimer (PDB id: 5GQ1) and PV 2C dimer (PDB id: 5Z3Q) found in crystal lattice. The adjacent 2C molecules are colored differently. (**B**, **D** and **F**) Magnified views at the 2C-2C juncture illustrating the PBD–Pocket interaction observed in the structures of HAV 2C, EV71 2C and PV 2C. The C-terminus helix is colored gold, and residues occupying the hydrophobic Pocket (highlighted with magenta surface) on are shown in a blue stick model.

tures and of homotypic interactions. These results also indicate that while ATPase domain, ZFER, PBD and Pocket of HAV 2C are all essential to virus replication, not all of them are required for the ATPase activity. The ZEFR of HAV 2C might be implicated in other vital functions of 2C yet to be identified. It is worth noting that the 2C mutants harboring an ATPase-dead mutations lacked the ATPase activity but still retained a minimal replication level compared to the 3D mut (a replication-deficient HAV replicon (33), Figure 4B–E, Supplementary Figure S4), suggesting that the ATPase function is important for viral replication but probably not strictly essential. However, ZFER, PBD and Pocket are involved in events - independent of ATPase activity—which are also important in virus replication.

HAV 2C exhibits ribonuclease activity that cleaves single-stranded RNA

Given that HAV 2C harbored all conserved SF3 helicase motifs, we characterized its unwinding activity. We used a helicase assay previously established for MERS-CoV NSP13 (30) based on an 18 bp partial RNA duplex with 10U 5′ overhang. Indeed, incubation with NSP13 resulted in an unwound product of the same size as the heat denatured control, indicating an authentic helicase activity (Figure 5A). In contrast, WT HAV 2C, as well as 2C ΔN, generated a reaction product neither matching in size to the dsRNA nor to the heated RNA (Figure 5A), suggesting this band was not the result of duplex unwinding. The same result was obtained with a substrate harboring a 3′ overhang and both reacting products had a size identical to the 18bp dsRNA portion of the substrate (Figure 5B). In line with the fact that the same activity was found for 2C WT and 2C ΔN, lacking parts of the ATP binding domain, the reaction proved to be independent of ATP in case of 2C WT (Supplementary Figure S7A). Thus, a plausible explanation was that HAV 2C digested the single stranded overhangs of both substrates instead of unwinding, suggesting a novel and unexpected nuclease activity. To rule out that the observed nuclease activity of HAV 2C was not a result of RNase contamination, we not only used Diethyl pyrocarbonate (DEPC) treated water for reaction buffer, but also included RNase inhibitor in all subsequent nuclease reactions. Comparing the nuclease reactions in the presence and absence of RNase inhibitor, we did not observe differences in RNA substrate digestion.To further understand substrate specificity of the potential nuclease activity, we analyzed HAV 2C digestion of different substrates. Indeed, 2C WT and ΔN both degraded a single stranded RNA substrate of 35nt in length, corresponding to one strand of the

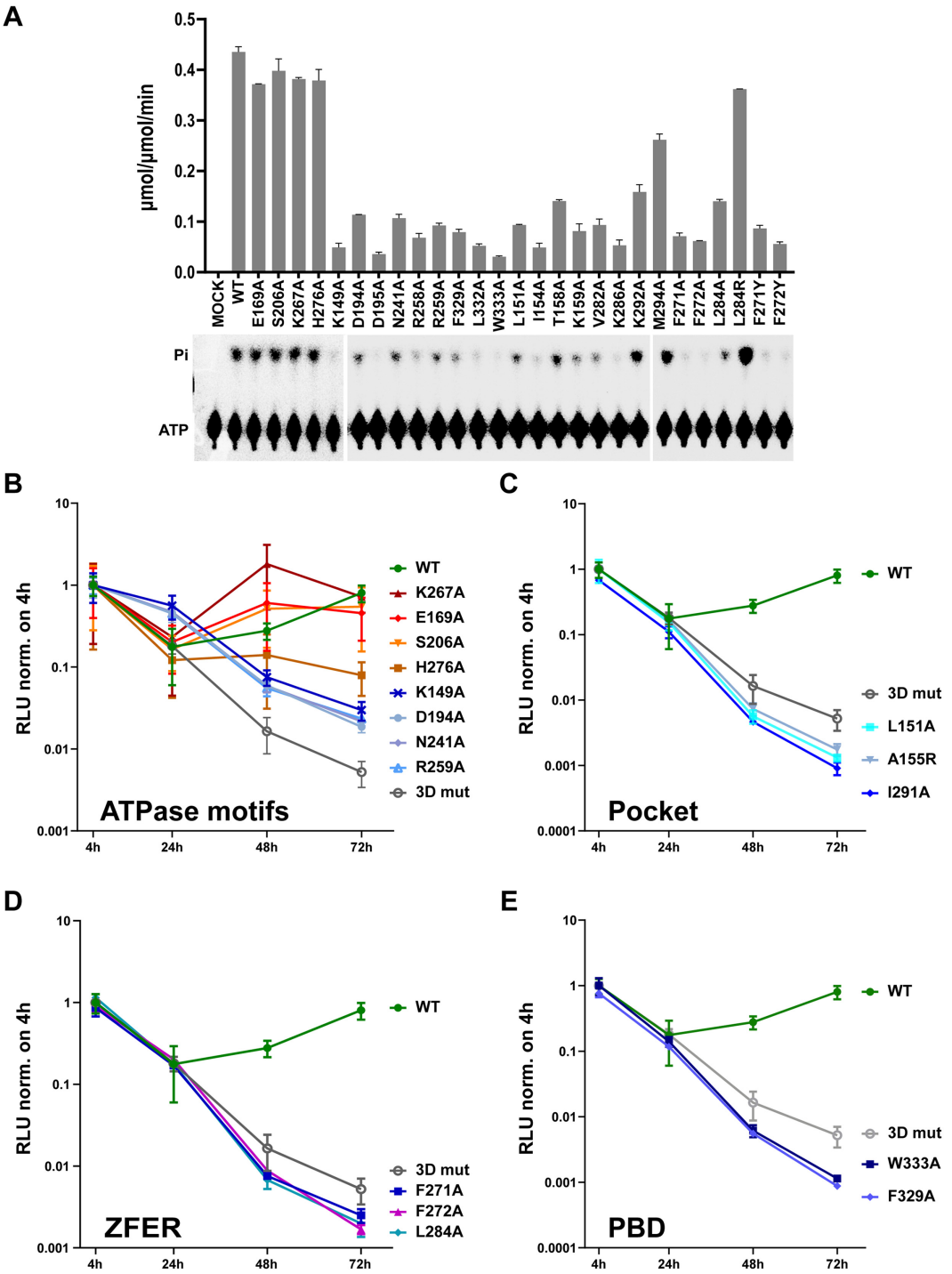

Figure 4. Impact of mutations on HAV 2C ATPase activity and viral RNA replication. (**A**) ATP activity of WT HAV 2C and a selection of 2C mutants. The ATPase activity is expressed as 'µmol of ATP hydrolyzed per µmol of enzyme per minute'. Images below are thin-layer chromatography of the ATP hydrolysis reaction visualized by Typhoon Trio Variable Mode Imager. (**B–E**). RNA replication efficiency of WT HAV and a selection of HAV 2C mutants. Huh7-Lunet cells were electroporated with subgenomic reporter replicons representing WT HAV or containing the indicated mutations in 2C region. At the respective timepoints, replication was determined by firefly luciferase assay in the lysates of transfected cells, representing replication efficiency. Data are normalized to the value obtained 4 hours after electroporation to correct for transfection efficiency. Mean values with SD from technical replicates of a representative experiment (n = 4 repetitions with comparable outcome). ATPase mutants and irrelevant mutants (refer to mutations outside of the conserved motifs critical to HAV 2C activities, E169A, S206A, K267A and H276A) are shown in panel B; Pocket (the hydrophobic pocket formed between ATPase domain and ZEFR) mutants are shown in panel C; ZFER (Zinc-finger equivalent region) mutants are shown in panel D; PBD (Pocket-Binding Domain) mutants are shown in panel E.

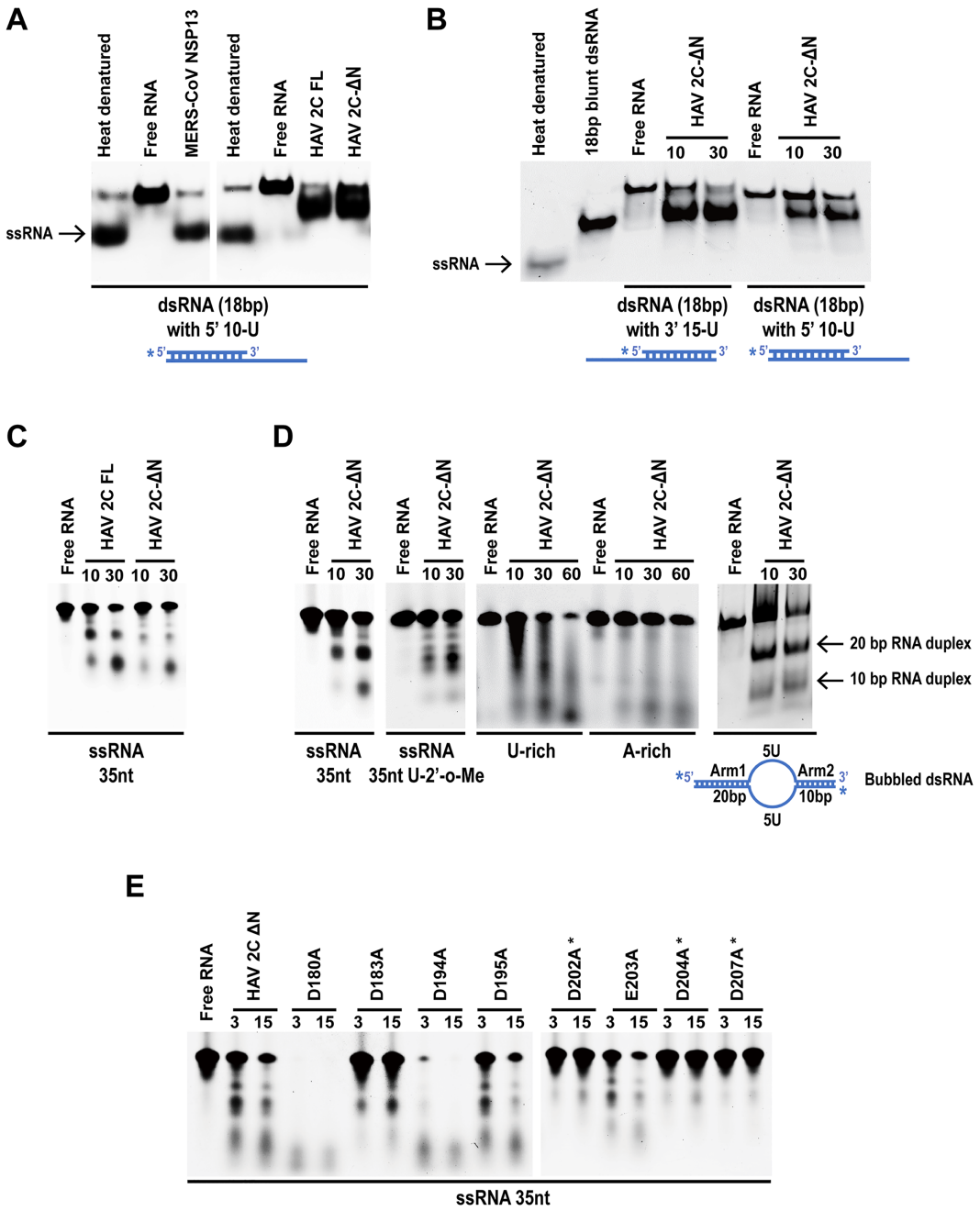

Figure 5. HAV 2C is a ribonuclease that specifically digests U-rich ssRNA regions. (**A**) Purified NSP13 of MERS-CoV, WT HAV 2C or 2C-ΔN were incubated with a dsRNA substrate containing a 18 bp duplex with 10U 5′ssRNA overhang (RNA1 & RNA4). The top strand of the duplex was labelled with Cy3, marked by an asterisk in the scheme below. The reaction mixtures were analysed by native-PAGE. Heat denatured substrate was used as a positive control for mimicking helicase activity. (**B**) HAV 2C-ΔN was incubated with two dsRNA substrates containing 18 bp duplex with 5′ (RNA1 & RNA4) or 3′ (RNA1 & RNA3) ssRNA overhangs. The top strand of the duplex was labelled with Cy3, marked by asterisks. The reaction mixtures were analysed by native-PAGE. 18bp dsRNA (RNA1 and RNA2) was loaded as a size marker. (**C**) Digestions of a Cy3 labelled 35 nt ssRNA by HAV 2C WT (FL) and 2C-ΔN. The incubation time is indicated above each lane and reaction mixtures were analyzed by urea–PAGE. (**D**) The same 35 nt ssRNA with 2′-O-Me modification at uridine residues (RNA6, left panel), a uridine-rich ssRNA (RNA7) and an adenosine-rich ssRNA (RNA8) were incubated with HAV 2C-ΔN and analysed as in (C) (middle panel). A dsRNA (RNA5 & RNA9) with an internal U-rich mismatch region flanked by 20 bp and 10 bp arms was used as a substrate and analyzed by Native-PAGE. (**E**) Digestion of a 35 nt ssRNA (RNA5) by HAV 2C-ΔN and a set of mutants harboring alanine substitution of the indicated acidic residues. The resulting mixtures were resolved by Urea-PAGE. Mutants lacking ribonuclease activity are marked by asterisk.

partial dsRNA substrate (Figure 5C). Since the truncated enzyme had comparable activity, we focused all subsequent experiments on this variant, due the far higher yields and purities obtained upon expression in *E. coli*, compared to the necessity to express and purify 2C WT from baculovirus infected insect cells. In agreement with the observed ssRNA specific RNase activity, we found that neither a dsRNA substrate nor a ssDNA were affected by treatment with 2C ΔN (Supplementary Figure S7B and C). Because methylated nucleotides confer nuclease resistance (40), we next analysed the 35-nt ssRNA containing 2′-O-methyluridine and indeed observed a different pattern, suggesting partial protection (Figure 5D). HAV 2C-ΔN furthermore digested a uridine-rich ssRNA more efficiently than an adenosine-rich ssRNA (Figure 5D, middle panels), suggesting that the enzyme prefers poly (U) stretches. Finally, we used a blunt-end double-strand RNA (dsRNA) containing a central mismatch bubble of five uridines, flanked by 20 bp and 10 bp duplex arms, respectively (Figure 5D, right panel). Incubation with HAV 2C-ΔN resulted in two major bands corresponding to dsRNA arms, suggesting that the enzyme was an endoribonuclease digesting the internal mismatch regions of the RNA without requiring a 5′ or 3′ terminus.

To identify critical residues responsible for the nuclease activity, we introduced a selection of mutations to HAV 2C-ΔN. We performed an alanine scan focusing on acidic residues, due to their importance to nuclease activity (40), including D180, D183, D194, D195, D202, E203, D204 and D207. Activity of these mutants was tested on the 35 nt ssRNA substrate (Figure 5E). Intriguingly, while D195A and E203A had little impact on the digestion of the RNA substrate, D180A and D194A exhibited dramatically increased nuclease activity. In stark contrast, the nuclease activity of mutant D183A was strongly impaired. Of note, individual alanine substitutions in an acidic cluster comprising D202, D204 and D207 all ablated RNase activity (Figure 5E), suggesting that this region is a part of nuclease active site. Using the nuclease-inactive mutant D207A, we measured RNA-binding affinity of HAV 2C-ΔN using biolayer interferometry (BLI). The binding dynamics between a biotinylated 35 nt ssRNA immobilized on streptavidin (SA) biosensor and HAV 2C-ΔN D207A supplied in serial dilutions from 35 to 2.2 μM was recorded, which gave a $K_d = 0.6$ μM (Supplementary Figure S7D). RNases usually binds their preferred nucleic acid substrates with high affinity. For example, *E. coli* RNase H1 digests RNA in RNA–DNA hybrid duplex, and it preferentially binds A-form duplex (RNA duplex or hybrid duplex, $K_d = 0.5$–3.4 μM) with much higher affinity than single-stranded nucleic acids ($K_d = 118$–942 μM) (41).

Collectively, our biochemical characterizations demonstrated that HAV 2C harbors, in addition to an ATPase activity, a previously unobserved ribonuclease activity specific on ssRNA regions of the substrate and with a preference to uridine stretches. Nuclease activity is independent of ATPase activity and involves an acidic cluster encompassing positions D202, D204 and D207.

Ribonuclease activity is conserved among picornavirus 2C proteins and potentially critical for RNA replication

We identified several acidic residues of HAV 2C critical for its ribonuclease activity. Since this region contains residues which are highly conserved among picornaviruses (Supplementary Figure S1) it seemed plausible to postulate that other 2C proteins possess the similar activities. To test this prediction, we prepared a set of N-terminal maltose-binding-protein (MBP) tagged full-length 2Cs from EV71, EV-D68, PV, FMDV, CVB3, Echo30 and HRV. All tested 2C proteins indeed showed similar levels of RNase activities, comparable to HAV2C on the 35 nt ssRNA substrate (RNA5) arguing for a conserved enzymatic function common to all picornaviral 2C proteins (Figure 6A). Surprisingly, none of the proteins apart from HAV was capable of cleaving the bubbled substrate RNA (Figure 6B), indicating subtle differences in the substrate specificity among picornavirus 2C variants. To further prove the specificity of the RNase activity residing in the picornaviral 2C proteins, we looked for invariant acidic residues critical for HAV 2C activity and thereby identified the aspartic acid corresponding to D204 (Figure 6C). We constructed the respective mutant for all picornaviral 2C proteins (FMDV 2C D170A, EchoV30 2C D186A; PV2C D186A; HRV 2C D179A; EV71 2C D186A; EV D68 2C D186A; CVB3 2C D186A) and tested the purified mutant proteins (Figure 6E) for RNase activity using the 35nt ssRNA substrate. As expected, all mutant proteins had lost the ability to degrade the substrate RNA (Figure 6D).

Finally, we used the subgenomic reporter replicon to investigate the role of ribonuclease activity of 2C in HAV replication, by introducing a set of mutations affecting its enzymatic function *in vitro* (Figure 7A, B). Intriguingly, all analyzed mutations significantly decreased virus replication, irrespective whether they increased or decreased ribonuclease activity (Figure 6A and B), arguing for an important function of this cluster of acidic residues. Importantly, no significant replication was found for those mutants entirely lacking nuclease activity (D202A, D204A, D207A, Figure 6C). To substantiate the importance of the RNase activity for other picornaviruses, we introduced the mutation D186A into an infectious cDNA clone of EV71 and compared the effect on virus production to a mutation in the ATPase active site (the Walker A), K135A, (Figure 7C) that abrogates EV71 replication (22). Both mutations resulted in the absence of detectable infectious virus, suggesting that the nuclease activity residing in 2C might serve important functions in the entire *Picornaviridae* family.

Taken together, our mutagenesis study indicates that the nuclease activity of HAV 2C might have an important function in HAV replication and also for other picornaviruses.

Hexameric-ring model of HAV 2C

We used the crystal structure of a hexameric SF3 helicase, the John Cunningham virus (JCV) large T antigen (PDB: 5J4Y), as the template for building a hexameric model of HAV 2C. After superimposing six copies of HAV 2C-ΔN onto each subunit of JCV large T antigen, we found that α6B helix harboring PBD clashed with the neighboring molecule (Figure 8A, left). This suggests that the C-terminal helix bend is an unnatural conformation, which was caused by crystal packing. Therefore, we adjusted the orientation of α6B to form a single, contiguous helix with the upstream α6A helix. In this conformation, the PBD could fit into the Pocket.

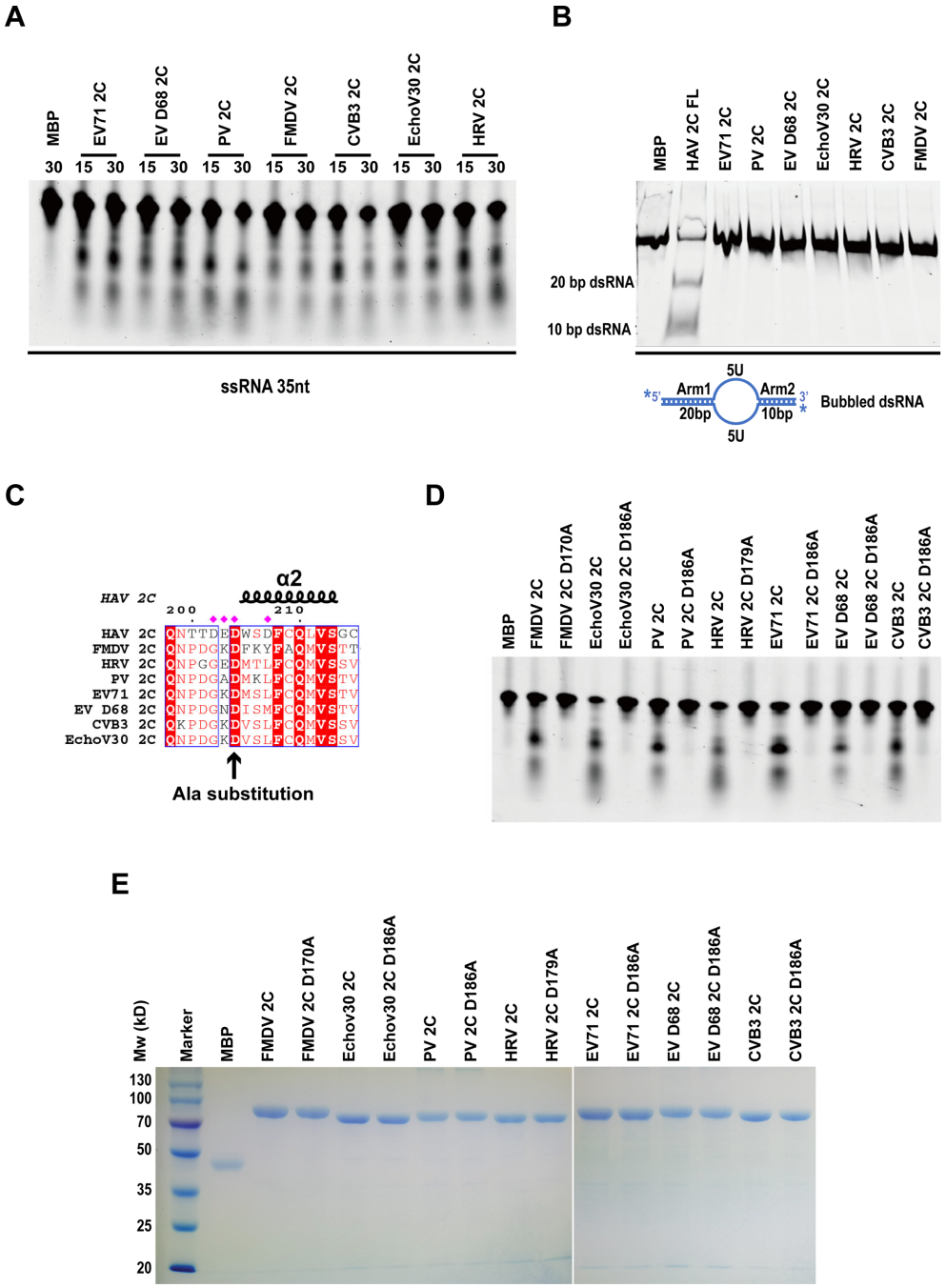

Figure 6. Multiple picornavirus 2C proteins exhibit ribonuclease activity. MBP-tagged full-length purified 2C proteins from Enterovirus 71 (EV71 2C), Enterovirus D68 (EV-D68 2C), Poliovirus (PV 2C), Foot-and-mouth disease virus (FMDV 2C), Coxsackievirus B3 (CVB3 2C), Echovirus 30 (EchoV30 2C) and Human Rhinovirus (HRV 2C), either WT or mutant, subjected to RNase assays using different substrates. (**A**) Picornaviral 2C proteins were incubated with a 35-nt single-stranded RNA substrate (RNA5) and the products were analyzed by 15% urea-PAGE. The incubation time (in minutes) is indicated on the top of the gel. (**B**) A bubble RNA substrate was prepared by annealing RNA5 with RNA9, and the substrate was incubated with the indicated 2C proteins at 37°C for 30 min. HAV 2C FL was the non-tagged WT protein. The products of the RNase assays were analyzed by 15% native PAGE. 'MBP' is the control of the MBP-tag alone. (**C**) An acidic residue D204 essential to HAV 2C RNase activity is invariant among various picornaviral 2C proteins (indicated by an arrow). (**D**) Various MBP-tagged 2C proteins and their mutants harboring alanine substitution of the invariant aspartic acid shown in panel C were tested for RNase activity on a 35-nt single-stranded RNA substrate (RNA5); the reaction mixtures were incubated at 37°C for 30 min before analysing by a 15% urea-PAGE. (**E**) SDS-PAGE analysis of various 2C proteins and their mutants used in the RNase assay shown in panel D.

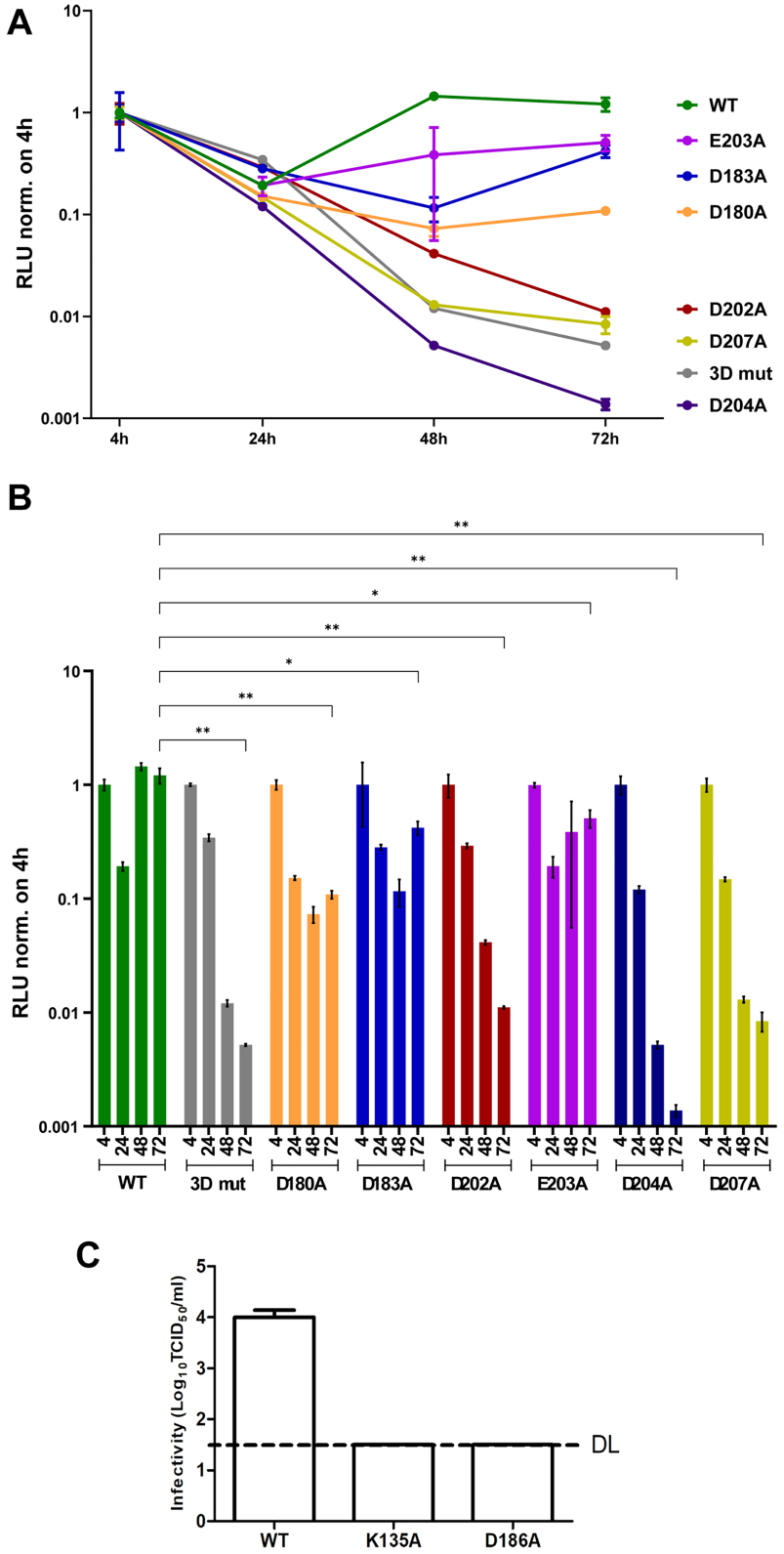

Figure 7. Impact of mutations in the RNase domain of HAV 2C and EV71 on replication. (**A**) Huh7-Lunet cells were electroporated with subgenomic HAV luciferase reporter replicon, wild-type (WT), replication dead (3D mut) or with mutations relevant to nuclease activity. At the indicated time points luciferase activity was determined, normalized to the 4-h value, and depicted in logarithmic scale. Graph depicts mean and standard deviation of two biological replicates performed in technical duplicates. RLU = relative light units. (**B**) Bar graph focusing on specific data of panel A. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$. (**C**) An infectious cDNA clone of EV71 was mutated at position D186A in the RNase domain (corresponding to D204A in HAV). A lethal mutation in the ATPase domain K135A (22) was used as a control. *In vitro* transcripts from WT and the respective mutants were transfected in Vero cells. Serial dilutions of freeze-thaw lysates obtained at 72 h after transfection were titrated on RD cells to quantify infectivity ($TCID_{50}$/ml). DL: detection limit.

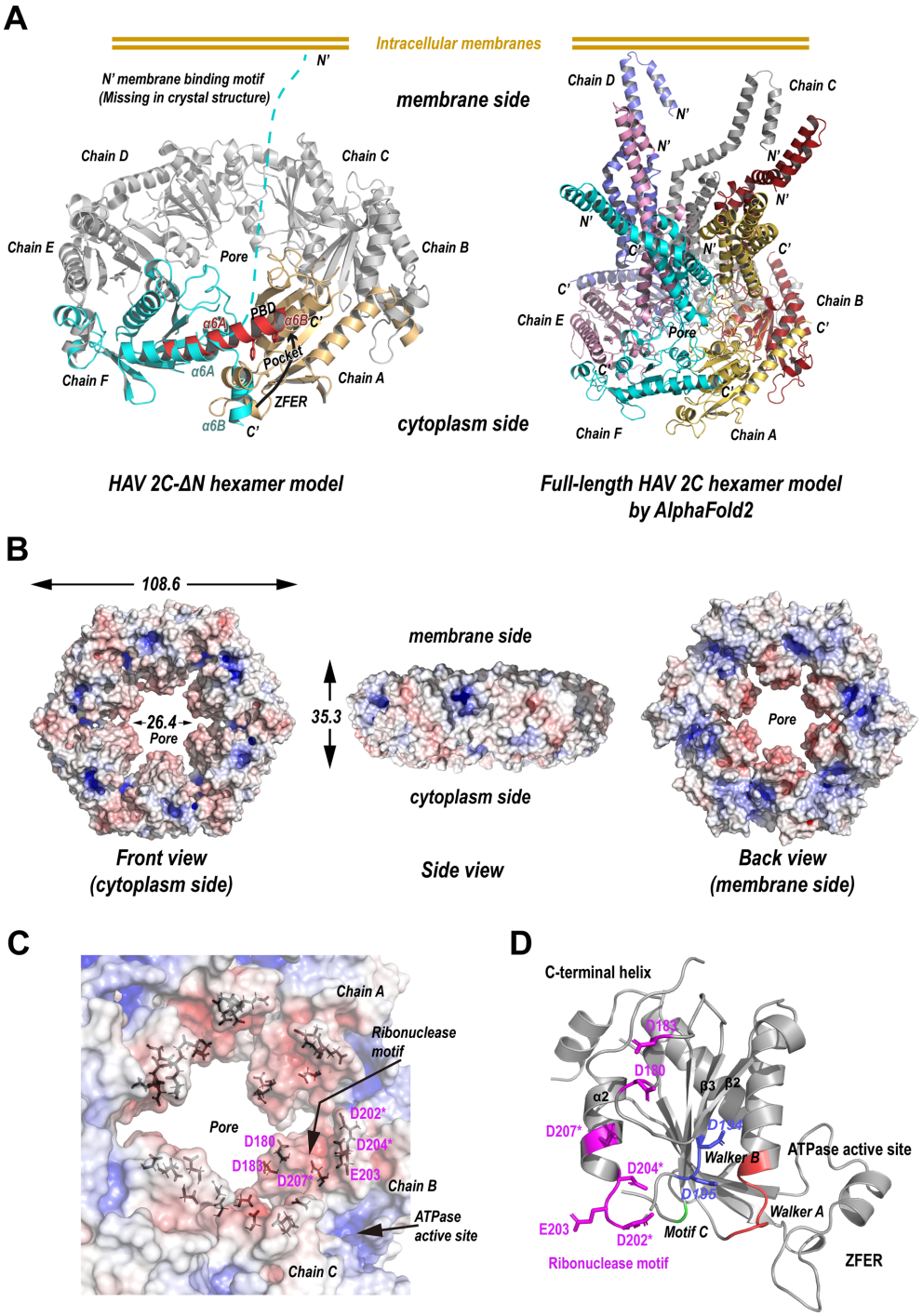

Figure 8. Hexameric model of HAV 2C. (**A**) Left, hexameric model of HAV 2C-ΔN was built by superimposing six copies of HAV 2C-ΔN structure to each subunit of JCV LTag hexamer (PDB: 5J4Y); the C-terminus α6B helix was merged with α6A helix, so that the PBD could be readily fitted into the Pocket on the adjacent 2C. The missing N-terminal portion of HAV 2C (Chain F) harboring membrane binding motif is indicated with the dashed line. Right, hexameric model of full-length HAV 2C. The monomeric model of full-length HAV 2C was predicted using program AlphaFold2. Six copies of full-length HAV 2C model were then superimposed to each subunit of the hexameric model of HAV 2C-ΔN. (**B**) Front, side and back views of surface electrostatic potential plot of the hexameric model of HAV 2C-ΔN. The dimensions of the hexameric model are marked. (**C**) Left, magnified view of the central pore on cytoplasm side; acidic residues important to ribonuclease activity (black stick model) are gathered around the pore on the cytoplasm side. Residues essential to nuclease activity of HAV 2C are indicated with asterisk. (**D**) Acidic residues important to ribonuclease activity are highlighted with stick model and colored magenta, Walker A motif is colored red, Walker B motif is colored blue. Residues essential to nuclease activity of HAV 2C are indicated with an asterisk.

The hexameric model of HAV 2C features a flat ring with a diameter ∼108.6 Å and a height ∼35.3 Å (Figure 8B), similar to the size of other 2C hexamers (22,23). The central pore of HAV 2C hexamer is unusually large (∼26.4 Å). This is attributed to the missing of α2-to-β4 loop (B′ motif) in our crystal structure, which is expected to line up at the pore. The ATPase active sites are located at the gaps between adjacent subunits, adopting reasonable geometries. While Walker A & B motifs and motif C lined up on one side of the gap, the R-finger was located on the other side. The closest distance between the R-finger (R259) and the Walker A motif (G146) in the hexameric model (∼4.2 Å) was significantly shorter than that in our crystal structure (∼30 Å) (Supplementary Figure S5).

We mapped the acidic residues important for nuclease activity of HAV 2C onto the hexameric model. We found all acidic residues gathered around the central pore on the cytoplasmic side of the ring (Figure 8C). Those residues do not overlap with the ATPase active sites (Figure 8D). Although the ATPase active sites in our hexameric model are separate from those RNase important residues, we cannot rule out possible links between the ATPase active sites and the putative RNase active sites; for example, large interdomain rearrangement may take place during ATP hydrolysis.

DISCUSSION

One main obstacle in understanding the function of picornavirus 2C is that we lack structural details of the N-terminal portion that harbors the putative membrane-binding motif. We previously determined two enterovirus 2C structures, but none included structural details for the N-terminal portion (the region preceding the ATPase domain) because in each case inclusion of the N-terminal region interfered with either solubility or crystallization. Therefore, we used the program AlphaFold2 (42) to predict the full-length HAV 2C structure (Supplementary Figure S6, Figure 8A, right). Our predicted models have high confidence score for the C-terminal domain (from 128 aa to end, CTD) and low model confidence at the N-terminal domain (1–127 aa, NTD), probably because there are no homologous structures available for HAV 2C NTD. We analyzed model 1 because it had the highest average pLDDT = 86.42 (Supplementary Figure S6A). Model 1 suggests that HAV 2C NTD and CTD are connected by a flexible loop (Supplementary Figure S6B). While the predicted CTD (average pLDDT = 92.3 indicating very high model confidence) matches well with our crystal structure, RMSD ∼1.4 Å, the predicted NTD had an average pLDDT = 76.7, suggesting moderate model confidence. The extreme N-terminal region of HAV 2C (1–37 aa) folds into two antiparallel amphipathic helices (Supplementary Figure S6B insert), which suggests that this region attaches to the surface of intracellular membranes rather than spanning membranes. Finally, we superimposed six copies of HAV 2C (from model 1) onto our crystallography-based HAV 2C-ΔN hexameric model (Figure 8A right) to build a comprehensive model. This representation suggests that the HAV 2C hexamer attaches to the surface of intracellular membrane with all six N-terminal amphipathic helices simultaneously, resembling a hexapod standing on a surface. Since AlphaFold2 can be unreliable for predicting the flexible linker between CTD and NTD, we must experimentally validate this hypothesis.

Picornavirus 2C is essential for virus replication, but its precise role in it remains to be understood. Common features shared by picornavirus 2Cs include the ability to hydrolyze NTPs and to bind intracellular membranes and RNAs. Banerjee et al. reported that the 2Cs of PV, HRV and HAV all bind to the 3′-UTR of the negative-strand viral RNA; thus, they postulated that 2C could immobilize negative-strand viral RNA to the membrane-anchored replication complex where it serves as the template for synthesizing plus-strand RNAs. Here, we provide evidence that HAV 2C harbors an unusual ssRNA-specific ribonuclease activity with a preference for poly (U) regions, similar to vertebrate ribonuclease RNase 4 (43). The substrate specificity and structural basis for the nuclease activity of other 2C proteins will require detailed investigations.

RNase activity of picornavirus 2C has remained undetected despite decades of research, probably because HAV 2C specifically binds the 3′ UTR of the negative-strand viral RNA, which lacks long poly(U) stretches and which folds into highly ordered structures that resist nucleases. On the other hand, synthesis of the positive-strand RNA requires priming by $VPgpUpU_{OH}$, a protein primer produced via CRE-dependent VPg uridylylation (44) for positive strand RNA synthesis, and probably directly on the polyA sequence for initiation of negative strand RNA synthesis (45). Therefore, it is tempting to speculate that the poly (U) targeting ribonuclease activity of HAV 2C might be implicated in the initial steps of positive- or negative strand RNA synthesis, e.g. the processing of VPg uridylylation products. The requirement of 2C for VPg uridylylation is evidenced by several studies demonstrating that guanidine hydrochloride, a reversible 2C inhibitor, can inhibit the CRE-dependent VPg uridylylation during PV replication (45,46). Alternatively, the 2C endonuclease might be involved in trimming of single stranded U-overhangs in double stranded replication intermediates or in cleaving off the VPg from the positive or negative strand genome, with yet to be defined functional significance. Another interesting question is why only HAV has retained the ability to cleave ssRNA in loop regions. Given that HAV is described as one of the most conserved primordial picornaviruses (47), it appears likely that this function might have been ancestral for 2C-associated nuclease activities. However, due to the abundance of functionally essential stem-loop structures in the genome of picornaviruses (48–51), this activity would endanger the integrity of the viral genome, since a single cleavage event could be deleterious. Therefore, the ssRNA cleavage might have got lost during the evolution of *Picornaviridae*, in which HAV diverged from the other members of the family. Still, this cleavage events might facilitate recombination, which has been proposed for HAV (52), within certain hotspot regions (53). All these potential functions are indeed highly speculative and hard to address experimentally, but certainly are worth pursuing.

Our conclusion that the RNase activity of 2C is important for viral RNA replication is currently based on the correlation of loss of function mutations in RNase activity and in the cell based assay. Of course, we cannot rule out that abrogation of RNA replication could be caused

by additional functions of the acidic residues beyond nuclease activity. Albeit further studies are required to define the precise function of the 2C ribonuclease activity in picornaviral RNA synthesis, it is to our knowledge the first viral RNase that appears to be crucial for the replication of an RNA virus identified so far. This activity is not related to E^{RNS} of pestiviruses, a secreted glycoprotein, that has been shown to degrade extracellular viral RNA contributing to innate immunotolerance (54). While we cannot exclude a function a function of the 2C RNase activity on cellular RNA, e.g. in counteracting innate immune responses, it appears rather unlikely for several reasons. First, the RNase activity of the purified protein is rather low, in line with the potential to endanger the integrity of the RNA genome. Second, the function is essential for RNA replication even in Huh7 cells, lacking a strong antiviral response (55), rather suggesting a function intrinsic to the viral replicase. Third, 2C typically is buried in the replicase complex, consisting of several viral non-structural proteins and strongly associated with virus induced membrane alterations (56), potentially precluding a widespread interaction with host RNAs.

HAV 2C exhibits both ATPase and RNA-binding activities; however, the helicase activity of HAV 2C has not been detected. The ssRNA-specific nuclease activity of HAV 2C identified in this study could have been confused with the unwinding activity; fortunately, this was clarified by our observation of the ATPase-inactive HAV 2C-ΔN yielding a similar result (Figure 5A). The absence of HAV 2C helicase activity is consistent with findings of many picornavirus 2Cs characterized to date, except EV71 2C (20). Comparing crystal structures of EV71 2C with PV 2C and HAV 2C, we could not identify a structural difference to explain their difference in helicase activity. We probably must determine structures of EV71 2C in complex with RNA duplex to reveal the mechanism of unwinding. Nevertheless, one question remains: If 2Cs do not hydrolyze ATP to unwind dsRNA, then why is 2C ATPase activity required for virus replication?

Combining our structural characterizations of multiple picornavirus 2Cs, we show that HAV 2C is not as distantly related to enterovirus 2Cs as suggested by sequence alignment. Apart from the conserved ATPase domains that fold into nearly identical structures in all available 2C structures, HAV 2C also undergoes C-terminal amphipathic helix (contact site 2) mediated self-oligomerization. We demonstrate that the PBD–Pocket interaction between HAV 2C molecules resembles the interaction between enterovirus 2C molecules. Despite being the most distant homolog of all 2Cs in *Picornaviridae*, HAV 2C retains this C-terminal-helix mediated self-oligomerization, suggesting that this mechanism is probably another feature (besides ATPase activity) conserved among picornavirus 2Cs. Further structural evidence from other 2Cs is necessary to support this postulation. On the other hand, the most structurally variable picornavirus 2C region lies between the ATPase domain and the C-terminal helix. Whereas this region in enterovirus 2C folds into a zinc-finger, the HAV 2C ZFER lacks 11 residues that include all zinc-coordinating cysteines. Of note, other 2Cs like FMDV 2C have a zinc-finger equivalent region of standard length, which also lacks cysteines or histidines to coordinate zinc. Further structural characterization is required in order to understand the function of the 2C ZFERs.

Design of anti-picornavirus drug focused on 2C because it is essential for picornavirus life-cycle. Numerous 2C inhibitors ranging from guanidine hydrochloride (57,58), HBB (59,60), MRL-1237 (61), TBZE-029 (62), fluoxetine (63–65), dibucaine (66), pirlindole (66) to zuclopenthixol (66) are available. Recently, Hurdiss and colleagues reported structures of CVB3 2C complexed by (S)-fluoxetine (SFX) (67). While SFX occupies a conserved pocket (SFX pocket) in the ATPase domain of CVB3 2C, a nearby '$_{224}$-AGSINA-$_{229}$' loop (the β4-to-β5 loop) specifically interact with the SFX pocket to govern its accessibility. Since the SFX pocket overlaps with the putative ribonuclease activity site that we identified in HAV 2C (Supplementary Figure S8), we measured the inhibitory effect of SFX on HAV 2C nuclease activity. First, we mapped key structural determinants for SFX binding. The HAV 2C SFX pocket resides in the cavity formed by α2, β2 and β3, adopting a 'closed' conformation because it is narrower than the CVB3 2C SFX pocket (Supplementary Figure S8A). Most CVB3 2C residues recognizing SFX are conserved in HAV 2C but the '$_{224}$-AGSINA-$_{229}$' loop of CVB3 2C is replaced by the entirely different '$_{243}$-WSNPSP-$_{248}$' loop in HAV 2C, challenging the accessibility of the pocket. Indeed, we detected no effect of SFX on HAV 2C nuclease activity, which supports the theory that the crosstalk between the SFX pocket and the β4-to-β5 loop is crucial for SFX efficacy.

By determining the crystal structure of EV71 2C and revealing the structural basis for PBD–Pocket dependent 2C oligomerization, we suggest a novel antiviral strategy that targets the Pocket for disrupting 2C-2C interaction (22). The PBD–Pocket interaction falls into the 'protein-protein interactions (PPIs)' drug-target category. Although PPIs were once considered 'undruggable', many clinically successful peptide-based PPI inhibitors have been designed on the basis of crystal structures of protein-protein complexes. Fang *et al.* designed and evaluated of a potent anti-enterovirus peptide derived from the C-terminal PBD of enterovirus 2C (68), confirming that the PBD–Pocket interaction can be targeted by a peptide-based inhibitor. Nonetheless, several knowledge gaps remain. (1) Does the PBD-derived peptide directly bind the 2C-pocket? If so, what is the binding affinity? This information is vital for validating the target and mechanism of the peptide and for additional inhibitor design. (2) Does the peptide inhibit 2C ATPase activity? Considering several picornavirus 2C crystal structures, we suggest the PBD–Pocket interaction could be probably druggable for a variety of picornaviruses. Not all picornavirus 2Cs are helicases but all are ATPases. Therefore, it is important to establish an ATPase-based inhibition assay to elucidate the mechanism of the inhibitor and to assess inhibitor efficacy *in vitro*.

DATA AVAILABILITY

Accession number for atomic coordinates: 7XT3 (Protein Data Bank).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Hongjie Zhang (Core Facility for Protein Research, Institute of Biophysics, Chinese Academy of Sciences) for the help with autoradiography. We thank the staff of PX III beamline at the Swiss Light Source, Paul Scherrer Institute (Villigen Switzerland) for assistance in data collection. We thank Rahel Klein (Molecular Virology, Center for Integrative Infectious Disease Research, University of Heidelberg, Germany) for excellent technical assistance and Yuri Kusov (Luebeck, Germany) for the generous gift of the HAV replicon. We thank the staffs of BL19U1 beamline of National Facility for Protein Science in Shanghai (NFPS) and the staff of the BL17U1 and BL02U1 beamline at the Shanghai Synchrotron Radiation Facility for assistance in data collection.

Author contributions: Sheng Cui and Pu Chen designed the study and wrote the paper. Sheng Cui, Justyna Aleksandra Wojdyla and Pu Chen solved the structure. Pu Chen, Ombretta Colasanti, Bo Qin and Zhijian Li performed experiments. Pu Chen, Justyna Aleksandra Wojdyla, Ombretta Colasanti, Zhijian Li, Meitian Wang, Volker Lohmann and Sheng Cui analyzed the data and revised the paper. All authors reviewed the results and approved the final version of the manuscript.

FUNDING

Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences [2021-I2M-1-043]; Collaborative Research Programme of International Centre for Genetic Engineering and Biotechnology [CRP/CHN19-02]; National Natural Science Foundation of China [82072291, 81772207, 81572005]; Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [272983813 – TRR 179, TP 17 to V.L.]. Funding for open access charge: International Centre for Genetic Engineering and Biotechnology.

Conflict of interest statement. None declared.

REFERENCES

1. International Committee on Taxonomy of Viruses (1995) In: *Virus Taxonomy: Classification and Nomenclature of Viruses: Sixth Report of the International Committee on Taxonomy of Viruses*. Springer-Verlag Wien Austria, NY.
2. Lemon,S.M., Ott,J.J., Van Damme,P. and Shouval,D. (2017) Type a viral hepatitis: a summary and update on the molecular virology, epidemiology, pathogenesis and prevention. *J. Hepatol.*, **68**, 167–184.
3. Averhoff,F.M., Khudyakov,Y. and Nelson,N.P. (2018) 24 - Hepatitis A Vaccines. In: Plotkin,S.A., Orenstein,W.A., Offit,P.A. and Edwards,K.M. (eds). *Plotkin's Vaccines. 7th edn*. Elsevier, pp. 319–341.
4. Foster,M.A., Hofmeister,M.G., Kupronis,B.A., Lin,Y., Xia,G.L., Yin,S. and Teshale,E. (2019) Increase in hepatitis a virus infections - United States, 2013-2018. *MMWR Morb. Mortal. Wkly. Rep.*, **68**, 413–415.
5. Hu,X., Collier,M.G. and Xu,F. (2020) Hepatitis a outbreaks in developed countries: detection, control, and prevention. *Foodborne Pathog Dis*, **17**, 166–171.
6. Stuart,D.I., Ren,J., Wang,X., Rao,Z. and Fry,E.E. (2019) Hepatitis a virus capsid structure. *Cold Spring Harb Perspect Med.*, **9**, a031807.
7. Debing,Y., Neyts,J. and Thibaut,H.J. (2014) Molecular biology and inhibitors of hepatitis a virus. *Med. Res. Rev.*, **34**, 895–917.
8. Jia,X.Y., Summers,D.F. and Ehrenfeld,E. (1993) Primary cleavage of the HAV capsid protein precursor in the middle of the proposed 2A coding region. *Virology*, **193**, 515–519.
9. Kusov,Y. and Gauss-Muller,V. (1999) Improving proteolytic cleavage at the 3A/3B site of the hepatitis a virus polyprotein impairs processing and particle formation, and the impairment can be complemented in trans by 3AB and 3ABC. *J. Virol.*, **73**, 9867–9878.
10. Gosert,R., Egger,D. and Bienz,K. (2000) A cytopathic and a cell culture adapted hepatitis a virus strain differ in cell killing but not in intracellular membrane rearrangements. *Virology*, **266**, 157–169.
11. Kuechler,E., Seipelt,J., Liebig,H.-D. and Sommergruber,W. (2002) Picornavirus Proteinase-Mediated Shutoff of Host Cell Translation: Direct Cleavage of a Cellular Initiation Factor. In: Semler,B. and Wimmer,E. (eds). *Molecular Biology of Picornavirus*. Wiley, pp. 299–311.
12. Pinto,R.M., Aragones,L., Costafreda,M.I., Ribes,E. and Bosch,A. (2007) Codon usage and replicative strategies of hepatitis a virus. *Virus Res.*, **127**, 158–163.
13. Kusov,Y.Y., Probst,C., Jecht,M., Jost,P.D. and Gauss-Muller,V. (1998) Membrane association and RNA binding of recombinant hepatitis a virus protein 2C. *Arch. Virol.*, **143**, 931–944.
14. Chen,P., Li,Z. and Cui,S. (2021) Chapter Nine -Picornaviral 2C proteins: An unique ATPase family critical in virus replication. In: Cameron,C.E., Arnold,J.J. and Kaguni,L.S. (eds). *The Enzyme*. Elsevier, Vol. **49**, pp. 235–264.
15. Joshi,M.S., Walimbe,A.M. and Chitambar,S.D. (2008) Evaluation of genomic regions of hepatitis a virus for phylogenetic analysis: suitability of the 2C region for genotyping. *J. Virol. Methods*, **153**, 36–42.
16. Teterina,N.L., Bienz,K., Egger,D., Gorbalenya,A.E. and Ehrenfeld,E. (1997) Induction of intracellular membrane rearrangements by HAV proteins 2C and 2BC. *Virology*, **237**, 66–77.
17. Yang,Y., Liang,Y., Qu,L., Chen,Z., Yi,M., Li,K. and Lemon,S.M. (2007) Disruption of innate immunity due to mitochondrial targeting of a picornaviral protease precursor. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 7253–7258.
18. Egger,D. and Bienz,K. (2005) Intracellular location and translocation of silent and active poliovirus replication complexes. *J. Gen. Virol.*, **86**, 707–718.
19. Singleton,M.R., Dillingham,M.S. and Wigley,D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, **76**, 23–50.
20. Xia,H., Wang,P., Wang,G.C., Yang,J., Sun,X., Wu,W., Qiu,Y., Shu,T., Zhao,X., Yin,L. *et al.* (2015) Human enterovirus nonstructural protein 2CATPase functions as both an RNA helicase and ATP-Independent RNA chaperone. *PLoS Pathog.*, **11**, e1005067.
21. Banerjee,R. and Dasgupta,A. (2001) Interaction of picornavirus 2C polypeptide with the viral negative-strand RNA. *J. Gen. Virol.*, **82**, 2621–2627.
22. Guan,H., Tian,J., Qin,B., Wojdyla,J.A., Wang,B., Zhao,Z., Wang,M. and Cui,S. (2017) Crystal structure of 2C helicase from enterovirus 71. *Sci. Adv.*, **3**, e1602573.
23. Guan,H., Tian,J., Zhang,C., Qin,B. and Cui,S. (2018) Crystal structure of a soluble fragment of poliovirus 2CATPase. *PLoS Pathog.*, **14**, e1007304.
24. Sweeney,T.R., Cisnetto,V., Bose,D., Bailey,M., Wilson,J.R., Zhang,X., Belsham,G.J. and Curry,S. (2010) Foot-and-mouth disease virus 2C is a hexameric AAA+ protein with a coordinated ATP hydrolysis mechanism. *J. Biol. Chem.*, **285**, 24347–24359.
25. Weinert,T., Olieric,V., Waltersperger,S., Panepucci,E., Chen,L., Zhang,H., Zhou,D., Rose,J., Ebihara,A., Kuramitsu,S. *et al.* (2015) Fast native-SAD phasing for routine macromolecular structure determination. *Nat. Methods*, **12**, 131–133.
26. Sheldrick,G.M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 479–485.
27. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
28. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
29. Liebschner,D., Afonine,P.V., Baker,M.L., Bunkoczi,G., Chen,V.B., Croll,T.I., Hintze,B., Hung,L.W., Jain,S., McCoy,A.J. *et al.* (2019)

Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in phenix. *Acta Crystallogr. D Struct. Biol.*, **75**, 861–877.
30. Hao,W., Wojdyla,J.A., Zhao,R., Han,R., Das,R., Zlatev,I., Manoharan,M., Wang,M. and Cui,S. (2017) Crystal structure of middle east respiratory syndrome coronavirus helicase. *PLoS Pathog.*, **13**, e1006474.
31. Esser-Nobis,K., Harak,C., Schult,P., Kusov,Y. and Lohmann,V. (2015) Novel perspectives for hepatitis a virus therapy revealed by comparative analysis of hepatitis c virus and hepatitis a virus RNA replication. *Hepatology*, **62**, 397–408.
32. van den Hoff,M.J., Moorman,A.F. and Lamers,W.H. (1992) Electroporation in 'intracellular' buffer increases cell survival. *Nucleic Acids Res.*, **20**, 2902.
33. Gauss-Muller,V. and Kusov,Y.Y. (2002) Replication of a hepatitis a virus replicon detected by genetic recombination in vivo. *J. Gen. Virol.*, **83**, 2183–2192.
34. Rodriguez,P.L. and Carrasco,L. (1993) Poliovirus protein 2C has ATPase and GTPase activities. *J. Biol. Chem.*, **268**, 8105–8110.
35. Adams,P., Kandiah,E., Effantin,G., Steven,A.C. and Ehrenfeld,E. (2009) Poliovirus 2C protein forms homo-oligomeric structures required for ATPase activity. *J. Biol. Chem.*, **284**, 22012–22021.
36. Cohen,J.I., Ticehurst,J.R., Purcell,R.H., Buckler-White,A. and Baroudy,B.M. (1987) Complete nucleotide sequence of wild-type hepatitis a virus: comparison with different strains of hepatitis a virus and other picornaviruses. *J. Virol.*, **61**, 50–59.
37. Yoon-Robarts,M., Blouin,A.G., Bleker,S., Kleinschmidt,J.A., Aggarwal,A.K., Escalante,C.R. and Linden,R.M. (2004) Residues within the B' motif are critical for DNA binding by the superfamily 3 helicase rep40 of adeno-associated virus type 2. *J. Biol. Chem.*, **279**, 50472–50481.
38. Holm,L. (2020) DALI and the persistence of protein shape. *Protein Sci.*, **29**, 128–140.
39. Laufman,O., Perrino,J. and Andino,R. (2019) Viral generated inter-organelle contacts redirect lipid flux for genome replication. *Cell*, **178**, 275–289.
40. Minskaia,E., Hertzig,T., Gorbalenya,A.E., Campanacci,V., Cambillau,C., Canard,B. and Ziebuhr,J. (2006) Discovery of an RNA virus 3′→5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 5108–5113.
41. Lima,W.F. and Crooke,S.T. (1997) Binding affinity and specificity of escherichia coli RNase H1: impact on the kinetics of catalysis of antisense oligonucleotide-RNA hybrids. *Biochemistry*, **36**, 390–398.
42. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Zidek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with alphafold. *Nature*, **596**, 583–589.
43. Sorrentino,S. (2010) The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Lett.*, **584**, 2194–2200.
44. Rieder,E., Paul,A.V., Kim,D.W., van Boom,J.H. and Wimmer,E. (2000) Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylylation. *J. Virol.*, **74**, 10371–10380.
45. Murray,K.E. and Barton,D.J. (2003) Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. *J. Virol.*, **77**, 4739–4750.
46. Steil,B.P. and Barton,D.J. (2009) Conversion of VPg into VPgpUpUOH before and during poliovirus negative-strand RNA synthesis. *J. Virol.*, **83**, 12660–12670.
47. Wang,X., Ren,J., Gao,Q., Hu,Z., Sun,Y., Li,X., Rowlands,D.J., Yin,W., Wang,J., Stuart,D.I. *et al.* (2015) Hepatitis a virus and the origins of picornaviruses. *Nature*, **517**, 85–88.
48. Andino,R., Rieckhof,G.E., Achacoso,P.L. and Baltimore,D. (1993) Poliovirus RNA synthesis utilizes an RNP complex formed around the 5′-end of viral RNA. *EMBO J.*, **12**, 3587–3598.
49. Duke,G.M., Hoffman,M.A. and Palmenberg,A.C. (1992) Sequence and structural elements that contribute to efficient encephalomyocarditis virus RNA translation. *J. Virol.*, **66**, 1602–1609.
50. Le,S.Y., Chen,J.H., Sonenberg,N. and Maizel,J.V. Jr (1993) Conserved tertiary structural elements in the 5′ nontranslated region of cardiovirus, aphthovirus and hepatitis a virus RNAs. *Nucleic Acids Res.*, **21**, 2445–2451.
51. Nagashima,S., Sasaki,J. and Taniguchi,K. (2003) Functional analysis of the stem-loop structures at the 5′ end of the aichi virus genome. *Virology*, **313**, 56–65.
52. Belalov,I.S., Isaeva,O.V. and Lukashev,A.N. (2011) Recombination in hepatitis a virus: evidence for reproductive isolation of genotypes. *J. Gen. Virol.*, **92**, 860–872.
53. Costa-Mattioli,M., Ferre,V., Casane,D., Perez-Bercoff,R., Coste-Burel,M., Imbert-Marcille,B.M., Andre,E.C., Bressollette-Bodin,C., Billaudel,S. and Cristina,J. (2003) Evidence of recombination in natural populations of hepatitis a virus. *Virology*, **311**, 51–59.
54. Lussi,C. and Schweizer,M. (2016) What can pestiviral endonucleases teach us about innate immunotolerance?*Cytokine Growth Factor Rev.*, **29**, 53–62.
55. Li,K., Chen,Z., Kato,N., Gale,M. Jr and Lemon,S.M. (2005) Distinct poly(I-C) and virus-activated signaling pathways leading to interferon-beta production in hepatocytes. *J. Biol. Chem.*, **280**, 16739–16747.
56. Harak,C. and Lohmann,V. (2015) Ultrastructure of the replication sites of positive-strand RNA viruses. *Virology*, **479-480**, 418–433.
57. Sadeghipour,S., Bek,E.J. and McMinn,P.C. (2012) Selection and characterisation of guanidine-resistant mutants of human enterovirus 71. *Virus Res.*, **169**, 72–79.
58. Tolskaya,E.A., Romanova,L.I., Kolesnikova,M.S., Gmyl,A.P., Gorbalenya,A.E. and Agol,V.I. (1994) Genetic studies on the poliovirus 2C protein, an NTPase. A plausible mechanism of guanidine effect on the 2C function and evidence for the importance of 2C oligomerization. *J. Mol. Biol.*, **236**, 1310–1323.
59. Hadaschik,D., Klein,M., Zimmermann,H., Eggers,H.J. and Nelsen-Salz,B. (1999) Dependence of echovirus 9 on the enterovirus RNA replication inhibitor 2-(alpha-Hydroxybenzyl)-benzimidazole maps to nonstructural protein 2C. *J. Virol.*, **73**, 10536–10539.
60. Klein,M., Hadaschik,D., Zimmermann,H., Eggers,H.J. and Nelsen-Salz,B. (2000) The picornavirus replication inhibitors HBB and guanidine in the echovirus-9 system: the significance of viral protein 2C. *J. Gen. Virol.*, **81**, 895–901.
61. Shimizu,H., Agoh,M., Agoh,Y., Yoshida,H., Yoshii,K., Yoneyama,T., Hagiwara,A. and Miyamura,T. (2000) Mutations in the 2C region of poliovirus responsible for altered sensitivity to benzimidazole derivatives. *J. Virol.*, **74**, 4146–4154.
62. De Palma,A.M., Heggermont,W., Lanke,K., Coutard,B., Bergmann,M., Monforte,A.M., Canard,B., De Clercq,E., Chimirri,A., Purstinger,G. *et al.* (2008) The thiazolobenzimidazole TBZE-029 inhibits enterovirus replication by targeting a short region immediately downstream from motif c in the nonstructural protein 2C. *J. Virol.*, **82**, 4720–4730.
63. Manganaro,R., Zonsics,B., Bauer,L., Lorenzo Lopez,M., Donselaar,T., Zwaagstra,M., Saporito,F., Ferla,S., Strating,J., Coutard,B. *et al.* (2020) Synthesis and antiviral effect of novel fluoxetine analogues as enterovirus 2C inhibitors. *Antiviral Res.*, **178**, 104781.
64. Hurdiss,D.L., El Kazzi,P., Bauer,L., Papageorgiou,N., Ferron,F.P., Donselaar,T., van Vliet,A.L.W., Shamorkina,T.M., Snijder,J., Canard,B. *et al.* (2022) Fluoxetine targets an allosteric site in the enterovirus 2C AAA+ ATPase and stabilizes a ring-shaped hexameric complex. *Sci. Adv.*, **8**, eabj7615.
65. Bauer,L., Manganaro,R., Zonsics,B., Strating,J., El Kazzi,P., Lorenzo Lopez,M., Ulferts,R., van Hoey,C., Mate,M.J., Langer,T. *et al.* (2019) Fluoxetine inhibits enterovirus replication by targeting the viral 2C protein in a stereospecific manner. *ACS Infect. Dis.*, **5**, 1609–1623.
66. Ulferts,R., de Boer,S.M., van der Linden,L., Bauer,L., Lyoo,H.R., Mate,M.J., Lichiere,J., Canard,B., Lelieveld,D., Omta,W. *et al.* (2016) Screening of a library of FDA-Approved drugs identifies several enterovirus replication inhibitors that target viral protein 2C. *Antimicrob. Agents Chemother.*, **60**, 2627–2638.
67. Hurdiss,D.L., El Kazzi,P., Bauer,L., Papageorgiou,N., Ferron,F.P., Donselaar,T., van Vliet,A.L.W., Canard,B., Decroly,E., Brancale,A. *et al.* (2022) Fluoxetine targets an allosteric site in the enterovirus 2C AAA+ ATPase and stabilizes the hexameric complex. *Sci. Adv.*, **8**, eabj7615.
68. Fang,Y., Wang,C., Wang,C., Yang,R., Bai,P., Zhang,X.Y., Kong,J., Yin,L., Qiu,Y. and Zhou,X. (2021) Antiviral peptides targeting the helicase activity of enterovirus nonstructural protein 2C. *J. Virol.*, **95**, e02324-20.