Crystal Structure of Human Enterovirus 71 3C Protease

Sheng Cui[1]*†‡, Jing Wang[1]†, Tingting Fan[1], Bo Qin[1], Li Guo[1], Xiaobo Lei[1], Jianwei Wang[1], Meitian Wang[2]*‡ and Qi Jin[1]*‡

[1]*State Key Laboratory for Molecular Virology and Genetic Engineering, Institute of Pathogen Biology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, People's Republic of China*
[2]*Swiss Light Source at Paul Scherrer Institute, CH-5232 Villigen, Switzerland*

Human enterovirus 71 (EV71) is the major pathogen that causes hand, foot and mouth disease that particularly affects young children. Growing hand, foot and mouth disease outbreaks were observed worldwide in recent years and caused devastating losses both economically and politically. However, vaccines or effective drugs are unavailable to date. The genome of EV71 consists of a positive sense, single-stranded RNA of ~7400 bp, encoding a large precursor polyprotein that requires proteolytic processing to generate mature viral proteins. The proteolytic processing mainly depends on EV71 3C protease (3C^{pro}) that possesses both proteolysis and RNA binding activities, which enable the protease to perform multiple tasks in viral replication and pathogen–host interactions. The central roles played by EV71 3C^{pro} make it an appealing target for antiviral drug development. We determined the first crystal structure of EV71 3C^{pro} and analyzed its enzymatic activity. The crystal structure shows that EV71 3C^{pro} has a typical chymotrypsin-like fold that is common in picornaviral 3C^{pro}. Strikingly, we found an important surface loop, also denoted as β-ribbon, which adopts a novel open conformation in EV71 3C^{pro}. We identified two important residues located at the base of the β-ribbon, Gly123 and His133, which form hinges that govern the intrinsic flexibility of the ribbon. Structure-guided mutagenesis studies revealed that the hinge residues are important to EV71 3C^{pro} proteolytic activities. In summary, our work provides the first structural insight into EV71 3C^{pro}, including a mobile β-ribbon, which is relevant to the proteolytic mechanism. Our data also provides a framework for structure-guided inhibitor design against EV71 3C^{pro}.

© 2011 Elsevier Ltd. All rights reserved.

*Corresponding authors. E-mail addresses: cuisheng2007@yahoo.com.cn; meitian.wang@psi.ch; jinqi@ipbcams.ac.cn.
† S.C. and J.W. contributed equally to this work.
‡ S.C., M.W., and Q.J. contributed equally to this work and are co-senior authors of the paper.

Abbreviations used: EV71, human enterovirus 71; HFMD, hand, foot and mouth disease; 3C^{pro}, 3C protease; FMDV, foot-and-mouth disease virus; HAV, hepatitis A virus; PV, poliovirus; HRV, human rhinovirus; CVB, coxsackievirus B; ASU, asymmetric unit; SARS-CoV, severe acute respiratory syndrome-coronavirus; WT, wild-type; PDB, Protein Data Bank; SLS, Swiss Light Source.

Introduction

Human enterovirus 71 (EV71) is the primary pathogen of herpangina or exanthema, also known as hand, foot and mouth disease (HFMD). Children and infants are especially vulnerable due to their underdeveloped immunocompetence.[1] Generally, HFMD starts with fever, headache, or vomiting. Lesions and rashes may become evident on patient's hand, foot and mouth in the following days, which give rise to the name of the disease. In severe cases, EV71 infections may lead to neurological damages, including aseptic meningitis, encephalitis, acute flaccid paralysis, myocarditis, pulmonary edema and hemorrhage[2] with high incidence of fatality. In recent years, the increasing outbreaks of HFMD have posed the global public health threats. HFMD epidemics were observed in China, Malaysia, Singapore, Australia, the United States and Germany.[3–8] In China, a total number of 192,300 cases were reported from January to April, 2010, an increase of almost 40% up compared to the same period in 2009. From the total number of cases, 2119 cases were diagnosed as severe and 94 cases were fatal.[9] An unusual outbreak of HFMD happened in Fuyang city (mainland China) in 2008, one of the most densely populated regions, causing the panic and endangering the social stability in the region. The initial fatality rate of the Fuyang outbreak was 2.9%, and all fatal cases died of serious neurogenic pulmonary edema were due to EV71 infections.[10] Unfortunately, effective drug or vaccine against EV71 is currently unavailable.[11]

EV71 belongs to the genus *Enterovirus* of the Picornaviridae family. The genome of EV71 consists of a positive sense, single-stranded RNA of ~7400 bp in length,[12] encoding a large precursor polyprotein that requires proteolytic processing to produce viral structural and replication proteins. In picornaviruses, except the cleavage between VP1 and 2A (catalyzed by the 2A protease[13]) and the RNA-dependent cleavage between VP2 and VP4,[14] the 3C protease (3C^{pro}) is absolutely required for all other polyprotein processing during the virus replication. Weng *et al.* discovered that EV71 3C^{pro} can interfere with the polyadenylation of host cellular RNA by digesting the host protein CstF-64, a critical factor for 3' pre-mRNA processing.[15] The finding proposed a novel mechanism by which the picornavirus utilizes 3C^{pro} to impair host cellular functions.

Most picornaviral 3C proteases preferentially cleave polypeptide with Gln–Gly junctions, with the exception of foot-and-mouth disease virus (FMDV) 3C^{pro}.[16] This unusual member of the picornaviral 3C^{pro} cleaves peptide junctions of P1–Gln or P1–Glu followed by a large apolar amino acid, such as leucine.[17–19] Structural studies on hepatitis A virus (HAV) 3C^{pro}, human rhinovirus

(HRV) 3C^{pro}, poliovirus (PV) 3C^{pro}, FMDV 3C^{pro} and coxsackievirus B (CVB) 3C^{pro} have shown that this family of cysteine protease shares an overall fold with many serine proteases, such as chymotrypsin, α-lytic protease and so on.[17,20–25] Picornaviral 3C proteases possess Cys-His-Glu/Asp catalytic triads that share geometry similar with that of the Ser-His-Asp triads found in serine proteases. Surprisingly, in the structure of HAV 3C^{pro} (with the Cys-His-Asp triad), the aspartic acid is redirected away from the active site. The Tyr143 points toward the catalytic histidine and provides the electrostatic stabilization.[26] This observation suggested the possible dyad proteolysis mechanism adopted by HAV 3C^{pro}. However, recent findings showing that both FMDV 3C^{pro} and the 3C-like tobacco etch virus protease[27] possess the Cys-His-Asp triads with the geometry similar to that of the Ser-His-Asp triad of serine protease, which prompted a reassessment of whether a dyad model is indeed the mechanism used by these family of proteases. Sequence and structural alignments show that the catalytic important motif Gly-X-Cys/Ser-Gly-Gly, serving to position Cys/Ser for nucleophilic attack and to form the oxyanion hole, is structurally conserved in picornaviral 3C proteases and serine proteases. Interestingly, the flexible surface loop β-ribbon is present in all picornaviral 3C^{pro}, such as FMDV 3C^{pro}, PV 3C^{pro}, HRV 3C^{pro},CVB3 3C^{pro} and HAV 3C^{pro}, and also in several bacteria serine proteases, such as the α-lytic protease,[28] *Streptomyces griseus* protease A[29] and *S. griseus* protease B.[30] The β-ribbon, located between the βB2 and βC2 strands, plays an important role in recognizing the P2–P4 region of peptide substrates. Regardless of the differences in length (12–21 aa) and in sequence, the β-ribbons maintain a similar fold and a similar orientation with respect to the substrate binding groove in these proteases.[17,18]

Beside the protease activities, picornaviral 3C^{pro} also has RNA binding activity, implying its role in picornaviral RNA replication.[31–33] By fusing with 3D^{Pol}, the RNA-dependent RNA polymerase that contains a nuclear localization sequence,[34] 3C^{pro} can enter nuclei and participate in viral RNA replication. The precursor protein 3CD is a crucial component of the viral replication machinery.[35] The PV 3CD functions as a protease better than PV 3C^{pro}, but lacks polymerase activity.[36] The autocatalytic activity enables 3CD to generate 3C^{pro} and 3D^{Pol}. Collectively, the multifunctional 3C^{pro} is an important player in many key events on the life cycle of the picornavirus. Therefore, the enzyme is generally considered as an appealing target for antiviral drug design.

In this work, we determined the first crystal structure of EV71 3C^{pro} and analyzed its protease activities. We found that EV71 3C^{pro} maintains the typical chymotrypsin fold that is similar to that of other picornaviral 3C^{pro}. Surprisingly, we found that

the β-ribbon of EV71 3C^{pro} is oriented in an unexpected and previously unobserved open conformation. Our structural analysis revealed two hinge residues located at the base of the β-ribbon. Supported by the mutagenesis studies, we found that the hinge residues likely govern the intrinsic flexibility of the β-ribbon. The mobility of the β-ribbon is relevant to the EV71 3C^{pro} protease activity. Our results also provide a structural framework for structure-based inhibitor design against EV71 3C^{pro}.

Results

Crystallization and structure determination

Three recombinant EV71 3C^{pro} variants (EV71 isolate BJ/CHN/2008) were highly expressed in *Escherichia coli* with yields of ∼2–10 mg of protein per liter of LB culture: N-terminal 6× His-tagged 3C^{pro}, C-terminal 6× His-tagged 3C^{pro} and non-tagged 3C^{pro}. All EV71 3C^{pro} variants were purified and concentrated to ∼12 mg ml^{-1} and were subjected to crystallization trials, only the N-terminal 6× His 3C^{pro} produced protein crystals. The crystals belonged to space group $P2_1$, contained five copies of EV71 3C^{pro} monomer per asymmetric unit (ASU) and diffracted X-ray to a resolution limit of 3.0 Å. The initial phases were obtained by molecular replacement. The phases were further improved by noncrystallographic symmetry averaging and manual model building, especially in the region of 123–133 aa where the novel β-ribbon conformation was

Table 1. Crystallographic data and model statistics

	Native
Data collection	
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	54.667, 75.59, 126.95
α, β, γ (°)	90, 90.13, 90
Wavelength	1.00
Resolution (Å)	3.0
R_{sym}	7.7 (38.5)
$I/\sigma I$	12.18 (2.74)
Completeness (%)	98.0 (92.0)
Redundancy	2.85
Refinement	
Resolution range (Å)	41.5–2.99
No. of reflections	39,192
R_{work}/R_{free}	20.6/25.9
No. of atoms	
Protein	7225
B-factors average	
Protein	82.84
r.m.s.d.	
Bond lengths (Å)	0.008
Bond angles (°)	1.066

Values for the highest-resolution shell are shown in parentheses.

observed. The refined model of EV71 3C^{pro} has $R/R_{free} = 0.21/0.26$ with good stereochemistry (Table 1).

Overall structure

EV71 3C^{pro} is folded in two topologically equivalent antiparallel six-stranded β-barrel domains (Fig. 1a and b) and is related to other picornaviral 3C^{pro} structures.[17,20,22,37] A long shallow groove for substrate binding is located between the two β-barrel domains. EV71 3C^{pro} shares ≤50% amino acid sequence identity with other picornaviral 3C^{pro},[17,20,24,25,37] among which CVB3 3C^{pro} shares the highest sequence similarity. The structural superimpositions (TOP program) give the r.m.s.d. of 0.87–1.04 Å between EV71 3C^{pro} (183 aa) and CVB3 3C^{pro}, HRV 3C^{pro} and PV 3C^{pro} for 160 C$^{\alpha}$ atoms. EV71 3C^{pro} shares only 24.2% and 29.1% sequence similarities with HAV 3C^{pro} and FMDV 3C^{pro}. The structural alignments map 110 C$_{\alpha}$ atoms of EV71 3C^{pro} onto structurally equivalent C$_{\alpha}$ atoms in HAV with r.m.s.d. value of 1.32 Å and map 132 C$_{\alpha}$ atoms of EV71 3C^{pro} onto structurally equivalent C$_{\alpha}$ atoms in FMDV 3C^{pro} with r.m.s.d. value of 1.40 Å. These data confirm that EV71 3C^{pro} shares an overall structural similarity with other picornaviral 3C^{pro}. However, EV71 3C^{pro} exhibits significant structural differences in the region corresponding to the β-ribbon located between βB2 and βC2 (123–133 aa). In other picornaviral 3C^{pro} structures, the β-ribbon is located over the substrate binding cleft with its apical tip pointing toward the protease active site. We denote this β-ribbon conformation as "close conformation". Strikingly, we found that the β-ribbon is flipped open with its tip pointing away from the active site in EV71 3C^{pro}. We denote this conformation as "open conformation" (Figs. 1a and b and 4a and b). In detail, at residue Gly123, the peptide backbone turns in the direction away from the active site, whereas at the corresponding position in the other picornaviral 3C^{pro}, the peptide backbone turns toward the active site. The opened loop extends for 11 aa and folds back at residue His133. From His133 onward, the backbone runs again in the direction similar to that observed in other picornaviral 3C^{pro} protease structures (Fig. 4b). Structural alignments of the β-ribbon (123–133 aa) of EV71 3C^{pro} and the corresponding regions in other picornaviral 3C^{pro} give r.m.s.d. values of <2 Å for 11 C$_{\alpha}$ positions, indicating that the opened β-ribbon still maintains the fold similar to that of the closed β-ribbons. In other words, while the β-ribbon itself remains rigid, the dihedral angles of residues G123 and H133 that are located at the base of the ribbon change dramatically, suggesting that these residues may function as hinges for the open-and-close flipping motion of the β-ribbon. Thus, the nature of the peptide bond restraints of the hinge residues could be the determining factor for the intrinsic flexibility of the β-ribbon.
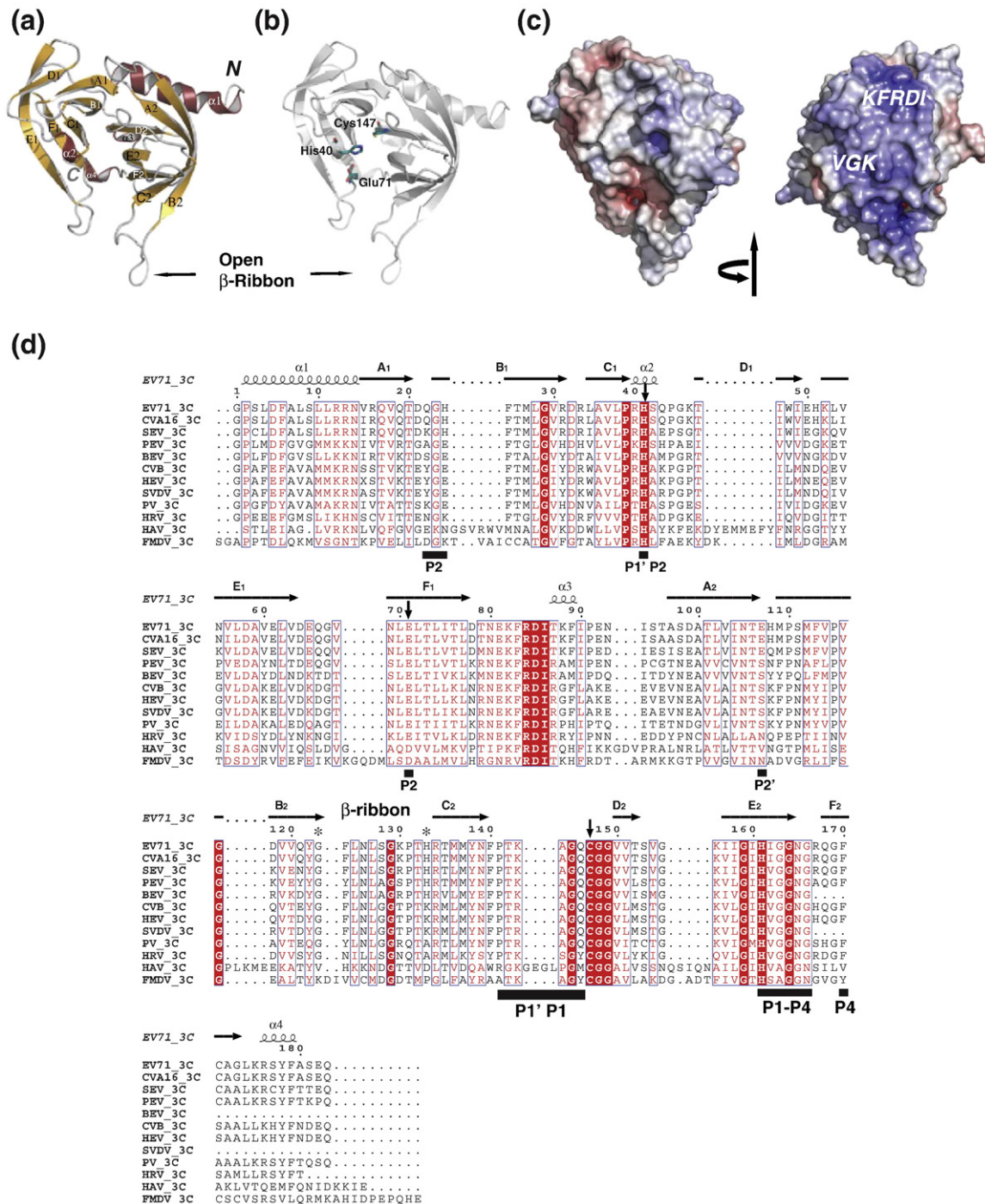
Fig. 1. Structure of EV71 3C[pro]. (a) Ribbon model of EV71 3C[pro] with annotated secondary structures. The position of β-ribbon is indicated with black arrow. (b) Ribbon model of EV71 3C[pro] (gray) with side chains of catalytic triad, His40, Glu71 and Cys147 (highlighted in stick presentation with deep teal for carbons). The position of β-ribbon is indicated. *There was no electron density for the side chain of residue Glu71. (c) Electrostatic surface potential (ranging from blue = 20 kT/e to red = −20 kT/e), displayed in two different views (left, "standard view" used in all other figures; right, 180° rotation around the vertical axis). Positions of the conserved RNA binding motifs "KFRDI" and "VGK" are indicated. (d) Structure-based multiple-sequence alignment of 3C[pro] of different picornaviruses. The secondary structure is shown on top. Invariant residues in 3C[pro] are highlighted with red background; conserved residues are shown in red font. The position of β-ribbon is indicated on the top of the sequences. Hinge residues Gly123 and His133 are indicated by asterisks. Black bars at the bottom of the sequence indicate residues directly involve in substrate binding (based on superimposition with HRV 3C–AG7088 co-structure; PDB code: 1CQQ).
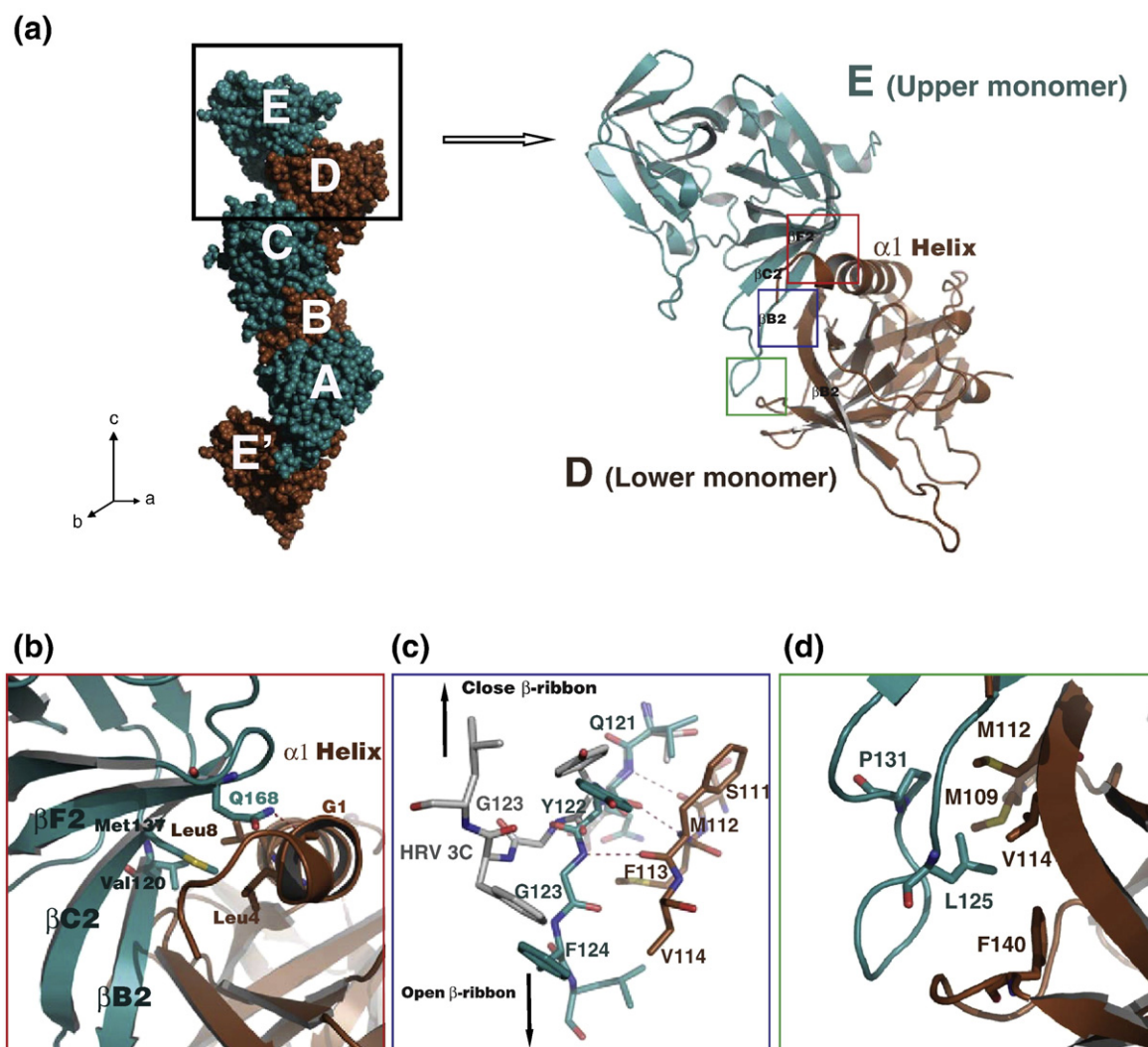
Fig. 2. Crystal packing stabilizes β-ribbon in open conformation. (a) Left side: five molecules are found in ASU: A, B, C, D and E. Molecule E′ is the corresponding E molecule from the unit cell below. EV71 3C$^{\text{pro}}$ monomers pack against each other, forming the continuous right-handed helical polymers extending alone the *c*-axis of the unit cell. The intermolecular interaction between any adjacent EV71 3C$^{\text{pro}}$ monomers is identical, which is magnified on the right side. Right side: ribbon model of molecule E/upper monomer (deep teal) and molecular D/lower monomer (brown), showing the intermolecular interaction between the EV71 3C$^{\text{pro}}$ monomers. Detailed crystal packing interactions are magnified in (b) (red box), (c) (blue box) and (d) (green box). (b) The N-terminal portion of the molecule D/lower monomer is packed against the surface of molecule E/upper monomer via hydrophobic interactions and hydrogen bond interaction (red broken line). (c) Stick representations showing the hydrogen-bonding interactions between the β-ribbon of molecule E/upper monomer (carbon atoms in deep teal) and the βB2 sheet of molecule D/lower monomer (carbon atoms in brown). Hydrogen bonds between the monomers are shown in red broken lines. The model of HRV 3C$^{\text{pro}}$ (carbon atoms in gray) is superimposed onto the molecule E/upper monomer. The directions to the tips of the open and close β-ribbons are indicated by arrows. (d) Leu125 and Pro131 located on the open β-ribbon of molecule E/upper monomer (deep teal) form the hydrophobic patch that interacts with the hydrophobic surface of molecule D/lower monomer (brown).

Crystal packing

We observed that the opened β-ribbon of EV71 3C$^{\text{pro}}$ is intensely involved in a particular crystal packing interaction. As shown in Fig. 2a (left), five EV71 3C$^{\text{pro}}$ monomers (molecules A–E) within the ASU pack against each other, forming a right-handed helical polymer with the helical axis parallel to the *c*-axis of the unit cell. Along the *c*-axis, the helical polymers of the adjacent unit cells are connected, forming the continuous helical polymer, in which the mode of the intermolecular interaction between any adjacent EV71 3C$^{\text{pro}}$ monomers is identical (Fig. 2a, right). In particular, on one side,

the N-terminal α1 helix of the lower monomer (molecule D) is packed against the concave surface formed by the βB2, βC2 and βF2 sheets of the upper monomer (molecule E). A hydrogen bond is found between Gly1 of the lower monomer and Gln168 of the upper monomer. Hydrophobic interactions are found between residue Leu4, residue Leu8 on the α1 helix of the lower monomer and residues Val120 and Met137 of the upper monomer (Fig. 2a and b). On the other side, the opened β-ribbon of the upper monomer (molecule E) is packed on the top of the C-terminal β-barrel domain of the lower monomer

(molecule D). In detail, three hydrogen bond interactions were observed between the β-ribbon of the upper monomer and the βB2 sheet of the lower monomer (Fig. 2a and c). The Gly123 of the upper EV71 3Cpro monomer is directed way from the position of its counterpart in HRV 3Cpro (Fig. 2c) by a hydrogen bond between the NH group of Gly123 and the carbonyl group of Phe113 of the lower EV71 3Cpro monomer. Two more intermolecular hydrogen bonds are found between Tyr122 of the upper monomer and Phe113 of the lower monomer and between Gln121 of the upper monomer and Ser111 of the lower monomer. These intermolecular hydrogen bonds contribute to maintain the β-ribbon of the upper EV71 3Cpro monomer in the open conformation. The open β-ribbon is further stabilized by hydrophobic interactions between the upper and lower monomers. Residues Leu125 and Pro131 located on the open β-ribbon of the upper EV71 3Cpro monomer form a hydrophobic patch that interact with the hydrophobic surface of the lower EV71 3Cpro monomer comprising of residues Met109, Met112, Val114 and Phe140 (Fig. 2d). In summary, the crystal packing analysis shows that the β-ribbon of the upper EV71 3Cpro monomer is tightly packed on top of the lower EV71 3Cpro monomer by means of a network of hydrogen bonds and hydrophobic interactions. These interactions contribute significantly to the stabilization of the open β-ribbon conformation.
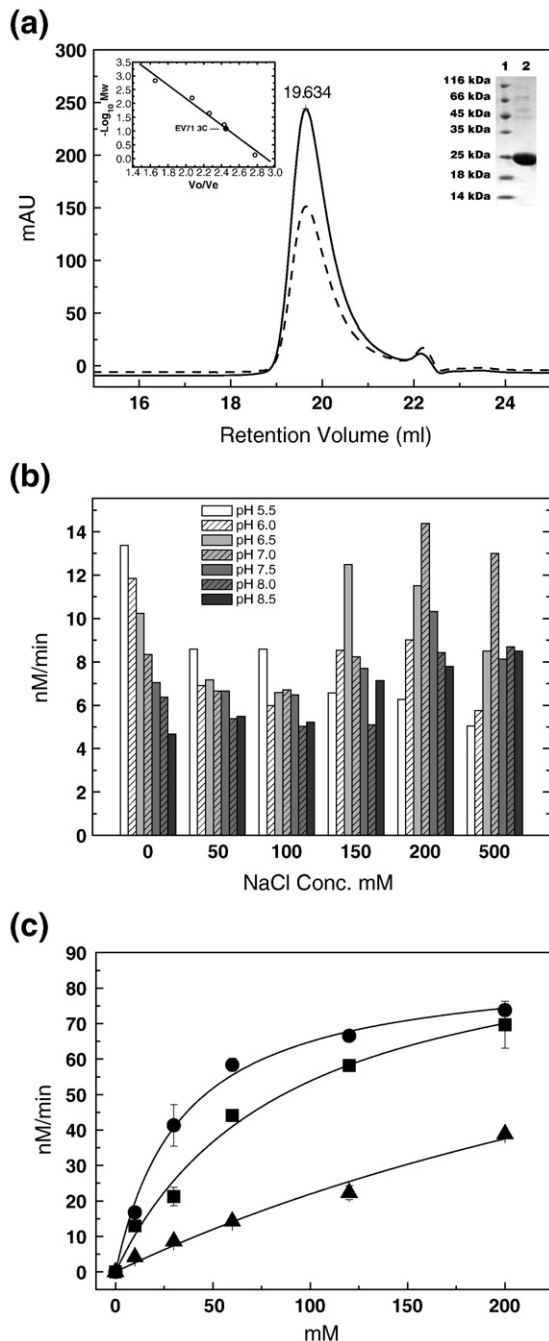
Structure of the active site

In the active site of EV71 3Cpro, the catalytic triads His40, Glu71 and Cys147 have a conformation similar to that observed in other picornaviral 3Cpro

Fig. 3. Biochemical characterization of EV71 3Cpro. (a) Size-exclusion chromatography and SDS-PAGE analysis of EV71 3Cpro protein. Chromatographic profiles are as follows: continuous line is UV absorbance at 280nm, and broken line is UV absorbance at 260nm. Upper-left inset: the −log (MW) values of standard proteins for size-exclusion column calibration (Superose 6 10/300GL) are plotted as the function of V_e/V_o (○). The data are fitted linearly to derive standard curve. The molecular mass of EV71 3Cpro is estimated at ∼17 kDa (●), indicating that EV71 3Cpro exists as monomer in solution. Upper-right inset: SDS-PAGE analysis of purified EV71 3Cpro. Lane 1, molecular mass standards; lane 2, EV71 3Cpro. (b) Optimization of protease assay conditions. The initial velocities of proteolytic reactions are measure in different reaction buffers with various combinations of salt (0–500 mM) and pH (pH 5.5–8.5). The best combination of salt and pH is determined as NaCl 200 mM (pH 7.0). (c) Substrate specificity of EV71 3Cpro. Fluorescence-quenching-pair-based proteolytic reactions are performed using different substrates: (●) substrate derived from the EV71 autoprocessing 3B–3C junction, (■) substrate derived from the first cleavage site (251/252 aa) in CstF-64 and (▲) substrate derived from the autoprocessing site of a 3C-like protease in SARS-CoV. The concentration of the protease was 4 μM for all experiments.

(Fig. 1b). However, the active site of EV71 3C^{pro} is very exposed to the solvent because the β-ribbon is in the open conformation. Poor electron density for the side chain of the catalytic important residue Glu71 did not allow for a definite model, and the residue is therefore associated with very high *B*-factor (an average *B*-factor of 121.59 Å^2 for Glu71 compare to an average *B*-factor of 82.84 Å^2 for all molecules in ASU). In order to find out whether or not EV71 3C^{pro} adopts the catalytic triad proteolytic mechanism, we performed mutagenesis studies. Mutants C147S, H40A and E71A were generated and subjected to protease assays. None of the mutants has detectable protease activities (data not shown), confirming that each member of the catalytic triad is essential for the protease activity of EV71 3C^{pro}. The catalytic important loop 141–147 aa is partially missing from the electron density map; thus, we could not observe the formation of the oxyanion hole, indicating that this region is disordered.

Structure of the RNA binding site

The KFRDI and VGK motifs were previously characterized as RNA binding motifs in picornaviral 3C^{pro}[25,37–39] and were confirmed as important residues for RNA binding in EV71 3C^{pro}.[31] In our EV71 3C^{pro} structure, the KFRDI motif is located in a loop connecting the βF1 and the α3 helix. The KFRDI motif forms a basic patch that sits at the opposite face with respect to the protease active site (Fig. 1c). The side chains of apolar residues Phe83 and Ile86 are inserted into the hydrophobic core of the protein, whereas the positively charged side chains of Lys82, Arg84 and Asp85 are exposed to the solvent. Besides Arg84 sitting at the center of the basic patch, the side chains of Arg12, Arg13 and Lys88 are also located at the positively charged surface. The VGK motif is located in the loop between strands βD2 and βE2. The side chain of Val156 is buried in the hydrophobic core, whereas Lys156 has its side chain pointing into the solvent. Based on our structure, more basic residues beside the KFRDI and VGK motifs are found in the putative RNA binding site of EV71 3C^{pro}, suggesting their involvement in RNA binding. The determination of EV71 3C^{pro}–RNA co-structure will be important in order to further elucidate the mode of EV71 3C^{pro}–RNA interaction.

Biochemical characterization of EV71 3C^{pro}

Size-exclusion chromatography analyses were carried out to determine the oligomerization state of the enzyme. The molecular mass of EV71 3C^{pro} determined by size-exclusion chromatography is 17 kDa (Fig. 3a, upper-left inset), which is similar to the molecular mass determined by SDS-PAGE (∼ 23 kDa; Fig. 3a, upper-right inset). The data suggest that EV71 3C^{pro} exists as a monomer in solution, supporting the notion that this class of proteases adopts a monomeric proteolysis mechanism.

In order to determine the optimal reaction conditions for EV71 3C^{pro} protease, we performed grid searches by varying salt concentrations and pH values. As shown in Fig. 3b, the combination of salt concentration and pH value that gives the best proteolytic activity is 200 mM NaCl and pH 7.0. Cations optimization was also performed by replacing sodium with potassium or by including magnesium, calcium, and so on, and only marginal differences were observed. Zinc was the only cation that negatively affected the protease activity of EV71 3C^{pro} (data not shown). This is likely due to the metal binding property of the catalytic Cys147. The optimized reaction buffer was used throughout the following experiments.

It has been reported that EV71 3C^{pro} cleaves most efficiently on the 3B–3C junction of EV71 polyprotein.[16] Recently, CstF-64 was identified as a novel cellular protein that serves as the target for EV71 3C^{pro}.[15] Therefore, we designed experiments to compare the enzyme activities of EV71 3C^{pro} against the enzyme activities of the EV71 polyprotein and the host target. The peptide corresponding to the EV71 polyprotein autoprocessing site 3B–3C (with the Q–G junction), the peptide corresponding to the internal cleavage site (251/252 aa with the Q–G junction) of CstF-64 and the peptide derived from the severe acute respiratory syndrome-coronavirus (SARS-CoV) autoprocessing site (with the Q–S junction) were used as the substrates in the protease assays. The protease concentration (4 μM) was kept constant in all protease assays. The kinetic parameters V_{max} and K_m were determined to calculate the specificity constant k_{cat}/K_m. As shown in Fig. 3c, EV71 3C^{pro} displayed the best enzyme activity against the autoprocessing 3B–3C site ($K_m = 30 \pm 2$ μM, $V_{max} = 85 \pm 1.4$ nM min^{-1} and $k_{cat}/K_m = 7.1 \times 10^{-4}$ μM^{-1} min^{-1}). Moderate enzyme activity was measured against the internal cleavage site of CstF-64 ($K_m = 87 \pm 15$ μM, $V_{max} = 91 \pm 7$ nM min^{-1} and $k_{cat}/K_m = 2.6 \times 10^{-4}$ μM^{-1} min^{-1}), and poor enzyme activity was measured in digesting the autoprocessing site of SARS-CoV (K_m and V_{max} could not be precisely determined by Michaelis–Menten equation fitting; thus, the k_{cat}/K_m was calculated from the linear part of the rate-*versus*-substrate concentration plot, $k_{cat}/K_m = 0.5 \pm 0.02 \times 10^{-4}$ μM^{-1} min^{-1}). The data suggest that EV71 3C^{pro} has substrate specificity and enzyme activity against the autoprocessing 3B–3C site greater than against the host target CstF-64 internal cleavage site. Although both substrates have Q–G junctions and P4–Ala residues, different residues at P2, P3, P4 and P2′ are also important for substrate recognition. The autoprocessing site of SARS-CoV with the Q–S junction is not the preferred substrate for EV71 3C^{pro}.
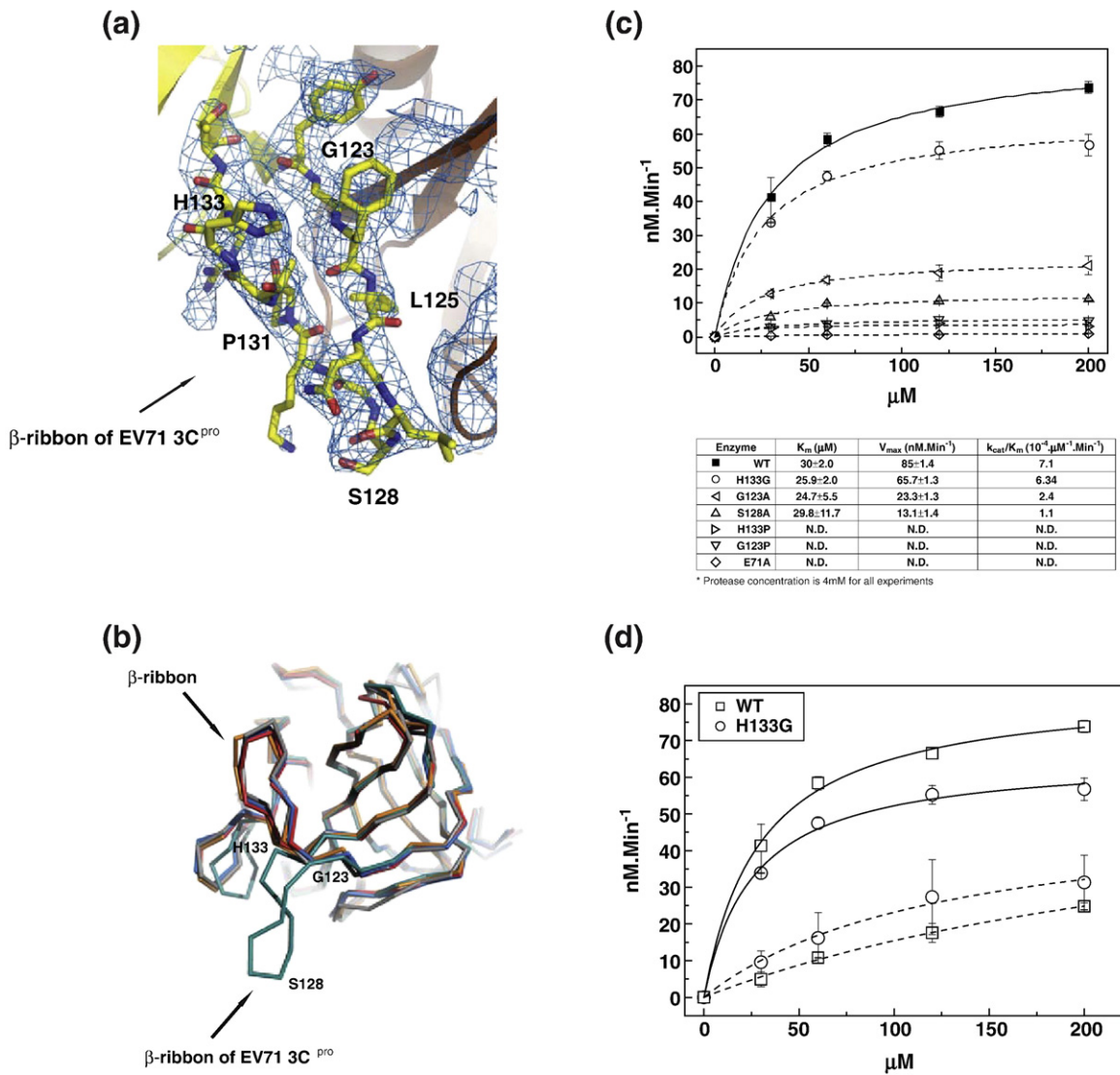
(a)

(c)

Enzyme		K_m (µM)	V_{max} (nM.Min^{-1})	k_{cat}/K_m (10^{-4} µM^{-1}.Min^{-1})
■	WT	30±2.0	85±1.4	7.1
○	H133G	25.9±2.0	65.7±1.3	6.34
◁	G123A	24.7±5.5	23.3±1.3	2.4
△	S128A	29.8±11.7	13.1±1.4	1.1
▷	H133P	N.D.	N.D.	N.D.
▽	G123P	N.D.	N.D.	N.D.
◇	E71A	N.D.	N.D.	N.D.

* Protease concentration is 4mM for all experiments

(b)

(d)

Fig. 4. The β-ribbon of EV71 3C^{pro}. (a) Portion of the final $2F_o - F_c$ electron density (1.0 s contour) at the opened β-ribbon with superimposed final model. (b) Ribbon model (backbone trace) superimpositions of EV71 3C^{pro} (deep-teal line) and other picornaviruses 3C^{pro}. Gray line, HRV 3C^{pro} (PDB code: 1CQQ); blue line, CVB3 3C^{pro} (PDB code: 2VB0); orange line, PV 3C^{pro} (PDB code: 1L1N); red line, CVB3 3C^{pro} (PDB code: 2ZTX). Positions of residues Ser128 sitting on the tip of the β-ribbon and the hinge residues Gly123 and His133 located at the base of the β-ribbon are indicated by the text. (c) Mutagenesis studies reveal the roles of residues in β-ribbon. Proteolytic reactions were performed using WT EV71 3C^{pro} and EV71 3C^{pro} mutants as the enzymes and fluorescence-labeled peptide derived from EV71 polyprotein 3B–3C junction as the substrate. Data for WT EV71 3C^{pro} are fitted with continuous line, and data for EV71 3C^{pro} mutants are fitted with broken lines. WT EV71 3C^{pro} (filled squares) serves as positive control; mutant E71A (open diamond) serves as non-active control. Mutants H133P (open right-pointing triangular) and G123P (open down-pointing triangular) are nearly inactive. Whereas, mutant H133G (open circle) is almost as active as WT protease. Mutant G123A (open left-pointing triangular) and mutant S128A (open up-pointing triangular) retain fractions of the enzyme activity. The table below summarizes the values of K_m and V_{max} and the ratio k_{cat}/K_m of the enzymes. (d) Comparison of substrate specificities of WT EV71 3C^{pro} and mutant H133G. Proteolytic reactions were performed using WT EV71 3C^{pro} (open square) and EV71 3C^{pro} mutant H133G (open circle) as the enzymes and fluorescence-labeled peptide derived from the EV71 polyprotein autoprocessing 3B–3C site (continuous lines) and peptide derived from the autoprocessing site in SARS-CoV (broken lines) as the substrates.

Mutagenesis studies revealed the mobile nature of the β-ribbon

The unusually open β-ribbon conformation is observed in EV71 3C^{pro} structure; however, in such case, the residues located on the opened β-ribbon cannot be positioned to form parts of S2–S4 substrate binding pockets. In HRV 3C^{pro}–AG7088 co-structure, the conserved Ser128 located on the tip of the β-ribbon is involved in substrate recognition by forming a hydrogen bond between its side-chain oxygen and backbone NH group of the inhibitor.[22] We found that EV71 3C^{pro} mutant S128A, in which the serine side-chain oxygen is missing, suffered from significant losses of the enzyme activity (Fig. 4c), confirming that the residue is important for substrate recognition. Therefore, we speculate that the EV71 3C^{pro} must also maintain the close β-ribbon conformation in solution. Supported by the structure of EV71 3C^{pro}, residues G123 and His133 located at the base of the β-ribbon may function as hinges. Thus, it is possible that the β-ribbon of EV71 3C^{pro} can undergo open and close motions around the hinges. To test this idea, we introduced a series of amino acid substitutions at the hinge positions. At one hinge, mutant G123P was inactive, whereas mutant G123A recovered a considerable fraction of activities. The K_m value for mutant G123A is similar to that of wild-type (WT) protease, but the V_{max} is reduced, indicating that the mutant has a similar substrate binding affinity, but suffers the losses of enzyme turnover rate. At the other hinge, no enzyme activity was observed for mutant H133P, whereas H133G had close-to-WT protease activity (Fig. 4c; K_m, V_{max} and k_{cat}/K_m values are listed in the table below). In general, when residues with more peptide bond restraints were introduced to the hinge positions, less protease activity was observed. When residues with less peptide bond restraints were introduced to the hinge positions, more protease activity was observed. In other words, the β-ribbon with the relaxed hinge favors the protease activity.

Subsequently, we compared the substrate specificities of mutant H133G (with the relaxed β-ribbon) and WT protease in processing different substrates. As shown in Fig. 4d, the WT protease cleaves the EV71 autoprocessing 3B–3C site with the specificity constant k_{cat}/K_m of 7.1×10^{-4} μM^{-1} min^{-1} ($K_m = 30 \pm 2$ μM; $V_{max} = 85 \pm 1$ nM min^{-1}); whereas the specificity constant for cleaving the autoprocessing site in SARS-CoV is $0.39 \pm 0.02 \times 10^{-4}$ μM^{-1} min^{-1}. Mutant H133G cleaves the autoprocessing 3B–3C site with the specificity constant k_{cat}/K_m of 6.34×10^{-4} μM^{-1} min^{-1} ($K_m = 25.9 \pm 2$ μM; $V_{max} = 65.7 \pm 1$ nM min^{-1}), and the mutant H133G cleaves the substrate of the autoprocessing site from SARS-CoV with the specificity constant of $0.69 \pm 0.04 \times 10^{-4}$ μM^{-1} min^{-1}. The data indicate that the substitution of His133 by glycine revealed no significant changes in substrate specificity.

Discussion

In the co-crystal structures of HRV 3C^{pro}–inhibitors, the β-ribbon adopts slightly different orientations when different inhibitors are bound.[37] Two very dissimilar β-ribbon conformations were found in different crystal forms of FMDV 3C^{pro}.[17,18] These studies indicate that the β-ribbon is intrinsically flexible. Our structural studies discovered a novel open conformation of the β-ribbon in EV71 3C^{pro}, providing a clear snapshot of the β-ribbon at the wide open extremity (Fig. 4a and b). The similar unusual β-ribbon conformation was first found in one of the crystal forms of FMDV 3C^{pro}, in which parts of the opened β-ribbon of FMDV 3C^{pro} were missing from the electron density map. It was proved that the unusual β-ribbon conformation of FMDV 3C^{pro} was due to the crystallization artifacts. We found that the β-ribbon of EV71 3C^{pro} is also involved in the extensive crystallographic packing interactions in the current crystal form. The crystal packing apparently contributes to the stabilization of the open β-ribbon conformation. Moreover, our structural studies demonstrated that although the orientation of the β-ribbon with respect to the body of the EV71 3C^{pro} changed dramatically, the β-ribbon itself still maintains a rigid fold. Importantly, two hinge residues were discovered at the base of the β-ribbon, which govern the flexibility of the β-ribbon. One of the hinge residue Gly123, an amino acid with minimal peptide bond restraints, is highly conserved (Fig. 1d), suggesting that the dynamic β-ribbon is a common feature of picornaviral 3C^{pro}. Residues such as Leu127 and Ser128 in HRV 3C^{pro} and Cys142 in FMDV 3C^{pro} located on the β-ribbon are usually important for substrate recognition (P2–P4). However, in the case of the open β-ribbon conformation, these residues are directed away from the substrate binding cleft and, thus, cannot participate in the substrate recognition. In the mutagenesis study, we determined that residue Ser128 located on the tip of the β-ribbon of EV71 3C^{pro} is involved in substrate recognition; however, the residue is located far away from the substrate binding cleft in the crystal structure of EV71 3C^{pro} due to the open β-ribbon conformation. Thus, the observed open β-ribbon conformation may only be transient, and the conformation was stabilized by crystallization artifact. We believe that EV71 3C^{pro} may adopt both open and close conformations. The protease may adopt the close β-ribbon conformation during substrate binding. Our mutagenesis studies also demonstrated that the increased mobility of the β-ribbon is correlated to the improved protease activity. However, both mutant with the restrained β-ribbon (G123A) and the mutant with relaxed β-ribbon (H133G) have the K_m value similar to that of WT protease, suggesting the similar substrate binding affinity. Thus, the data suggest that the

mobility of the β-ribbon may be important to the product production steps subsequent to the initial substrate binding.

In summary, our data suggest a scenario where the substrate recognition and hydrolysis involve the movement of the β-ribbon. The β-ribbon is in motion in the absence of the substrate. The motion mode of the β-ribbon can be described as a moving clamp that flips between the open and close positions around the hinge residues. The mobility of the β-ribbon is determined by the peptide bond restraints of the hinge residues. The transiently opened β-ribbon results in the more exposed substrate binding cleft, thus may increase the substrate accessibility. When the substrate is bound, the β-ribbon can interact with the N-terminal side of the substrate and then stabilizes at the close conformation, leading to the enzyme–substrate complex. The β-ribbon may adopt different orientations upon substrate binding, depending on the structure of the substrate. Subsequently, the mobility of the β-ribbon may be required for the efficient disassociation of the N-terminal of the cleaved substrate after the acylation half of the reaction.

AG7088 (Rupintrivir) is an HRV 3C^{Pro} inhibitor designed to combat common cold caused by rhinovirus infections. Interestingly, there are increasing evidences showing that AG7088 is also effective against enteroviruses.[16,40] In order to shed light on the structural basis of these observations, we modeled AG7088 into substrate binding cleft of EV71 3C^{Pro} by superimposing the EV71 3C^{Pro} structure onto the HRV 3C^{Pro}–AG7088 co-crystal structure (Fig. 5). The modeled AG7088 is fitted to the substrate binding cleft of EV71 3C^{Pro} similarly as observed in HRV 3C^{Pro}–AG7088 interactions, suggesting that HRV 3C^{Pro} and EV71 3C^{Pro} share the similar architecture of substrate binding cleft and similar mechanisms for substrate recognition.

EV71 3C^{Pro} has no known cellular homologue in humans. Marginal differences in the sequences of 3C^{Pro} proteins are observed among enteroviruses and coxsackieviruses causing HFMD.[40] This limited sequence diversity suggests the significant structural similarity, which may offer the possibility of targeting the protease with single-drug therapies. Recently, based on the architecture of AG7088, specific inhibitors against EV71 3C^{Pro} were designed.[16] Our structural insight into EV71 3C^{Pro} can certainly provide valuable information for further structure-based inhibitor developments.

Materials and Methods

Constructs and proteins preparation

The cDNA encoding the EV71 3C^{Pro} (1–183 aa) was amplified by RT-PCR experiment using the genomic RNA of EV71 (EV71 isolate BJ/CHN/2008; GenBank
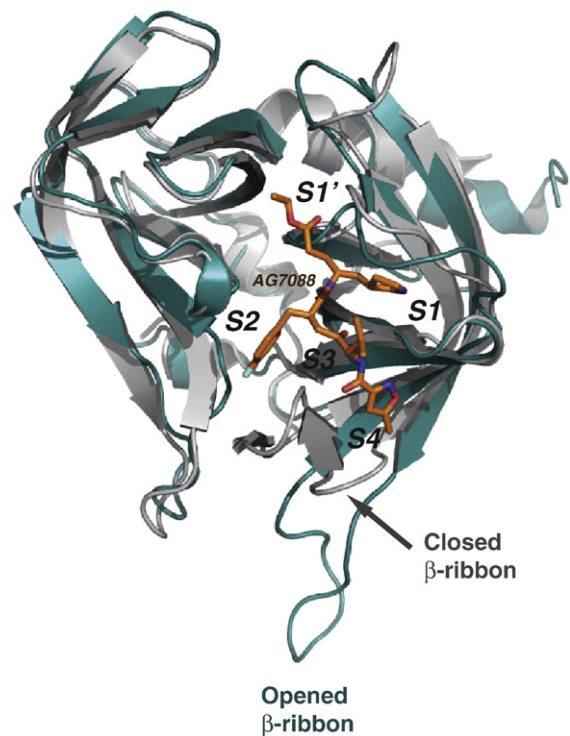

Fig. 5. Superimposition of EV71 3C^{Pro} and HRV 3C^{Pro}–AG7088 structures. Ribbon model of EV71 3C^{Pro} structure (deep teal) is superimposed onto the ribbon model of HRV 3C^{Pro}–AG7088 co-structure (PDB code: 1CQQ). HRV 3C^{Pro} is colored gray, and AG7088 is colored by elements (carbon atoms in orange). The substrate binding pockets, S1′–S4, are indicated. The opened β-ribbon of EV71 3C^{Pro} and the closed β-ribbon of HRV 3C^{Pro} are indicated.

accession number: HQ615421) as template. The sequence of the resulting cDNA was verified by DNA sequencing and was inserted into pET28a or pET21a vectors (Novagen) to generate plasmids expressing N-terminal 6× His-tagged, non-tagged and C-terminal 6× His-tagged EV71 3C^{Pro} variants. The plasmid expressing the N-terminal-tagged EV71 3C^{Pro} variant encodes full-length EV71 3C^{Pro} (183 aa) and the N-terminal extension: MGSSHHHHHHSSGLVPRGSHM, containing 6× His, a linker region and a thrombin cleavage site (underlined). The non-tagged EV71 3C^{Pro} variant was obtained by thrombin digestion followed by additional purification steps to remove thrombin contamination. The construct expressing the C-terminal EV71 3C^{Pro} variant encodes full-length EV71 3C^{Pro} plus the C-terminal extension: AAA-LEHHHHHH. Plasmids expressing EV71 3C^{Pro} mutants were generated using the site-directed mutagenesis methods (QuikChange™) and were verified by DNA sequencing. The plasmids were transformed into *E. coli* Rosetta competent cells (Novagen). The bacteria cultures were grown at 37 °C to an OD$_{600}$ = 1.0 and were cooled down to 18 °C before induction (0.5 mM IPTG). The bacteria cultures were shaken at 18 °C overnight and were subsequently harvested and disrupted by ultrasonication. All recombinant EV71 3C^{Pro} variants and mutants were purified to

homogeneity using multiple chromatographic steps following standard protocols. Briefly, batches of EV71 3C^{pro} lysate were initially purified using affinity chromatography (Ni-NTA; Qiagen) followed by cation-exchange chromatography (HighTrap SP; GE Healthcare). The final step of purification was size-exclusion chromatography (Sephacryl S-200 HR 26/60 column; GE Healthcare).

Size-exclusion experiments

Superose 6 10/300GL column (GE Healthcare) was pre-equilibrated with a buffer containing 50 mM Tris–HCl (pH 7.5), 150 mM NaCl and 1 mM DTT. The column was then calibrated using gel-filtration molecular mass standards (Bio-Rad), containing thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (45 kDa), myoglobin (17 kDa) and vitamin B$_{12}$ (1.35 kDa). The retention volumes of the standards were measured, and the values of V_e (retention volume)/V_o (void volume) were plotted as the function of $-\log_{10}(MW)$ to derive the standard curve. EV71 3C^{pro} was loaded on the column to determine V_e. The molecular weight of EV71 3C^{pro} was calculated using the standard curve.

Protease activity assay

The synthetic peptide substrates were attached with a fluorescence quenching pair, Dabcyl and Edans. The fluorogenic peptides Dabcyl-RTATV**QG**PSLDFE-Edans corresponding to the EV71 polyprotein autoprocessing site between 3B–3C, Dabcyl-MQASM**QG**GVPAPE-Edans derived from the cleavage site (251/252 aa) in CstF-64 and Dabcyl-TSAVL**QS**GFRKM-Edans derived from the N-terminal autoprocessing site of a 3C-like protease from SARS-CoV were used to measure the protease activities. Fluorescence experiments were performed with Spectra-Max® M5 Multi-Mode Microplate Reader. We measured 100 μl of reaction mixture containing 50 mM Tris–HCl (pH 7.0), 200 mM NaCl, 2 mM DTT and 4 μM EV71 3C^{pro} or its mutants and fluorogenic peptides of different concentrations in a 96-well plate (Greiner) at 30 °C. The instrument was pre-calibrated by Edans standard to calculate the relationship between relative fluorescence unit and concentration. The relative fluorescence unit was collected using an excitation wavelength of 340$_{nm}$ and by monitoring the emission 500$_{nm}$ and was then converted to substrate concentration. The initial velocities of the reactions are plotted *versus* substrate concentrations. Values are averaged from three independent experiments. The data were fitted to Michaelis–Menten equation to calculate K_m, V_{max} and k_{cat}/K_m (MicroCal Origin). For data that the K_m and V_{max} values cannot be precisely determined by Michaelis–Menten equation fitting, the k_{cat}/K_m was calculated from the linear part of the rate-*versus*-substrate concentration plot.

Crystallization and structure determination

The EV71 3C^{pro} variants were concentrated to ∼12 mg ml^{-1} prior to the crystallization trials, only N-terminal 6× His-tagged EV71 3C^{pro} resulted in useful crystals. The optimal crystallization of EV71 3C^{pro} was achieved by mixing 0.8 μl protein with 1.0 μl buffer containing 100 mM Tris–HCl (pH 7.7), 200 mM sodium citrate and 20% polyethylene glycol 3350 in a hanging-drop vapor diffusion system at 22 °C. Cryocooling was achieved by soaking the crystals for 30–60 s in reservoir solution containing 15% ethylene glycol and by flash-freezing them in liquid nitrogen. Complete data to a resolution limit of 3.0 Å on native crystals were collected at the PXIII beamline of the Swiss Light Source (SLS; Villigen, Switzerland). The data sets were processed with the XDS Package.[41] The space group was identified as $P2_1$, and five molecules were found in the ASU. The initial phases were obtained by the program Phaser v2.1,[42] using the structure of CVB3 3C^{pro} [Protein Data Bank (PDB) ID: 2VB0] structure as search model. The electron density map was further improved by non-crystallographic symmetry averaging. The manual model building was carried out with the program Coot v0.6.[43] Prior to model building and refinement, we randomly omitted 5% of the reflections for monitoring the R_{free} value. Refinement was performed with PHENIX[44] and included overall anisotropic B-factor and bulk solvent corrections, individual B-factor refinement, simulated annealing and positional refinement. Data collection and model statistics are summarized in Table 1. Figures were prepared with PyMOL (DeLano Scientific).

Accession number

The coding sequence of EV71 isolate BJ/CHN/2008 3C protein has been deposited in GenBank with the accession number HQ615421.

Coordinates and structure factors have been deposited in the PDB with accession number 3OSY.

References

1. Chen, C. Y., Chang, Y. C., Huang, C. C., Lui, C. C., Lee, K. W. & Huang, S. C. (2001). Acute flaccid paralysis in

infants and young children with enterovirus 71 infection: MR imaging findings and clinical correlates. *Am. J. Neuroradiol.* **22**, 200–205.

 2. Lee, T. C., Guo, H. R., Su, H. J., Yang, Y. C., Chang, H. L. & Chen, K. T. (2009). Diseases caused by enterovirus 71 infection. *Pediatr. Infect. Dis. J.* **28**, 904–910.

 3. McMinn, P., Stratov, I., Nagarajan, L. & Davis, S. (2001). Neurological manifestations of enterovirus 71 infection in children during an outbreak of hand, foot, and mouth disease in Western Australia. *Clin. Infect. Dis.* **32**, 236–242.

 4. Ho, M., Chen, E. R., Hsu, K. H., Twu, S. J., Chen, K. T., Tsai, S. F. *et al.* (1999). An epidemic of enterovirus 71 infection in Taiwan. Taiwan Enterovirus Epidemic Working Group. *N. Engl. J. Med.* **341**, 929–935.

 5. Zhang, S. B., Liao, H., Huang, C. H., Tan, Q. Y., Zhang, W. L., Huang, Y. *et al.* (2008). [Serum types of enterovirus and clinical characteristics of 237 children with hand, foot and mouth disease in Shenzhen]. *Zhongguo Dangdai ErKe ZaZhi*, **10**, 38–41.

 6. AbuBakar, S., Chee, H. Y., Al-Kobaisi, M. F., Xiaoshan, J., Chua, K. B. & Lam, S. K. (1999). Identification of enterovirus 71 isolates from an outbreak of hand, foot and mouth disease (HFMD) with fatal cases of encephalomyelitis in Malaysia. *Virus Res.* **61**, 1–9.

 7. Alexander, J. P., Jr., Baden, L., Pallansch, M. A. & Anderson, L. J. (1994). Enterovirus 71 infections and neurologic disease—United States, 1977–1991. *J. Infect. Dis.* **169**, 905–908.

 8. Kehle, J., Roth, B., Metzger, C., Pfitzner, A. & Enders, G. (2003). Molecular characterization of an Enterovirus 71 causing neurological disease in Germany. *J. NeuroVirol.* **9**, 126–128.

 9. Yan, Central China province reports surge in HFMD deaths http://news.xinhuanet.com/english, 2010.

10. The Chinese Center for Disease Control and Prevention, and the World Health Organization, Report on the Hand, Foot and Mouth Disease Outbreak in Fuyang City, Anhui Province and the Prevention and Control in China (May 2008) World Health Organization Representative Office in China, 2008.

11. Zhang, D., Lu, J. & Lu, J. (2010). Enterovirus 71 vaccine: close but still far. *Int. J. Infect. Dis.* **14**, e739–e743.

12. Yang, F., Jin, Q., He, Y., Li, L. & Hou, Y. (2001). The complete genome of Enterovirus 71 China strain. *Sci. China, Ser. C: Life Sci.* **44**, 178–183.

13. Hwang, Y. C., Chen, W. & Yates, M. V. (2006). Use of fluorescence resonance energy transfer for rapid detection of enteroviral infection *in vivo*. *Appl. Environ. Microbiol.* **72**, 3710–3715.

14. Bishop, N. E. & Anderson, D. A. (1993). RNA-dependent cleavage of VP0 capsid protein in provirions of hepatitis A virus. *Virology*, **197**, 616–623.

15. Weng, K. F., Li, M. L., Hung, C. T. & Shih, S. R. (2009). Enterovirus 71 3C protease cleaves a novel target CstF-64 and inhibits cellular polyadenylation. *PLoS Pathog.* **5**, e1000593.

16. Kuo, C. J., Shie, J. J., Fang, J. M., Yen, G. R., Hsu, J. T., Liu, H. G. *et al.* (2008). Design, synthesis, and evaluation of 3C protease inhibitors as anti-enterovirus 71 agents. *Bioorg. Med. Chem.* **16**, 7388–7398.

17. Birtley, J. R., Knox, S. R., Jaulent, A. M., Brick, P., Leatherbarrow, R. J. & Curry, S. (2005). Crystal structure of foot-and-mouth disease virus 3C protease.

18. Sweeney, T. R., Roqué-Rosell, N., Birtley, J. R., Leatherbarrow, R. J. & Curry, S. (2007). Structural and mutagenic analysis of foot-and-mouth disease virus 3C protease reveals the role of the beta-ribbon in proteolysis. *J. Virol.* **81**, 115–124.

19. Zunszain, P. A., Knox, S. R., Sweeney, T. R., Yang, J., Roqué-Rosell, N., Belsham, G. J. *et al.* (2009). Insights into cleavage specificity from the crystal structure of foot-and-mouth disease virus 3C protease complexed with a peptide substrate. *J. Mol. Biol.* **395**, 375–389.

20. Lee, C. C., Kuo, C. J., Ko, T. P., Hsu, M. F., Tsui, Y. C., Chang, S. C. *et al.* (2009). Structural basis of inhibition specificities of 3C and 3C-like proteases by zinc-coordinating and peptidomimetic compounds. *J. Biol. Chem.* **284**, 7646–7655.

21. Fuhrmann, C. N., Kelch, B. A., Ota, N. & Agard, D. A. (2004). The 0.83 Å resolution crystal structure of alpha-lytic protease reveals the detailed structure of the active site and identifies a source of conformational strain. *J. Mol. Biol.* **338**, 999–1013.

22. Matthews, D. A., Dragovich, P. S., Webber, S. E., Fuhrman, S. A., Patick, A. K., Zalman, L. S. *et al.* (1999). Structure-assisted design of mechanism-based irreversible inhibitors of human rhinovirus 3C protease with potent antiviral activity against multiple rhinovirus serotypes. *Proc. Natl Acad. Sci. USA*, **96**, 11000–11007.

23. Hopfner, K. P., Lang, A., Karcher, A., Sichler, K., Kopetzki, E., Brandstetter, H. *et al.* (1999). Coagulation factor IXa: the relaxed conformation of Tyr99 blocks substrate binding. *Structure*, **7**, 989–996.

24. Bergmann, E. M., Cherney, M. M., Mckendrick, J., Frormann, S., Luo, C., Malcolm, B. A. *et al.* (1999). Crystal structure of an inhibitor complex of the 3C proteinase from hepatitis A virus (HAV) and implications for the polyprotein processing in HAV. *Virology*, **265**, 153–163.

25. Mosimann, S. C., Cherney, M. M., Sia, S., Plotch, S. & James, M. N. (1997). Refined X-ray crystallographic structure of the poliovirus 3C gene product. *J. Mol. Biol.* **273**, 1032–1047.

26. Allaire, M., Chernaia, M. M., Malcolm, B. A. & James, M. N. (1994). Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. *Nature*, **369**, 72–76.

27. Phan, J., Zdanov, A., Evdokimov, A. G., Tropea, J. E., Peters, H. K., 3rd, Kapust, R. B. *et al.* (2002). Structural basis for the substrate specificity of tobacco etch virus protease. *J. Biol. Chem.* **277**, 50564–50572.

28. Fujinaga, M., Delbaere, L. T., Brayer, G. D. & James, M. N. (1985). Refined structure of alpha-lytic protease at 1.7 Å resolution. Analysis of hydrogen bonding and solvent structure. *J. Mol. Biol.* **184**, 479–502.

29. James, M. N., Sielecki, A. R., Brayer, G. D., Delbaere, L. T. & Bauer, C. A. (1980). Structures of product and inhibitor complexes of *Streptomyces griseus* protease A at 1.8 Å resolution. A model for serine protease catalysis. *J. Mol. Biol.* **144**, 43–88.

30. Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., Jr & James, M. N. (1982). Refined crystal structure of the molecular complex of *Streptomyces griseus* protease B, a serine protease, with the

third domain of the ovomucoid inhibitor from turkey. *Proc. Natl Acad. Sci. USA*, **79**, 4868–4872.

31. Shih, S. R., Chiang, C., Chen, T. C., Wu, C. N., Hsu, J. T., Lee, J. C. *et al.* (2004). Mutations at KFRDI and VGK domains of enterovirus 71 3C protease affect its RNA binding and proteolytic activities. *J. Biomed. Sci.* **11**, 239–248.

32. Peters, H., Kusov, Y. Y., Meyer, S., Benie, A. J., Bäuml, E., Wolff, M. *et al.* (2005). Hepatitis A virus proteinase 3C binding to viral RNA: correlation with substrate binding and enzyme dimerization. *Biochem. J.* **385**, 363–370.

33. Kusov, Y. Y. & Gauss-Muller, V. (1997). *In vitro* RNA binding of the hepatitis A virus proteinase 3C (HAV 3Cpro) to secondary structure elements within the 5′ terminus of the HAV genome. *RNA*, **3**, 291–302.

34. Andino, R., Rieckhof, G. E., Achacoso, P. L. & Baltimore, D. (1993). Poliovirus RNA synthesis utilizes an RNP complex formed around the 5′-end of viral RNA. *EMBO J.* **12**, 3587–3598.

35. Parsley, T. B., Towner, J. S., Blyn, L. B., Ehrenfeld, E. & Semler, B. L. (1997). Poly (rC) binding protein 2 forms a ternary complex with the 5′-terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA*, **3**, 1124–1134.

36. Harris, K. S., Reddigari, S. R., Nicklin, M. J., Hämmerle, T. & Wimmer, E. (1992). Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. *J. Virol.* **66**, 7481–7489.

37. Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W. *et al.* (1994). Structure of

human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. *Cell*, **77**, 761–771.

38. Leong, L. E., Walker, P. A. & Porter, A. G. (1993). Human rhinovirus-14 protease 3C (3Cpro) binds specifically to the 5′-noncoding region of the viral RNA. Evidence that 3Cpro has different domains for the RNA binding and proteolytic activities. *J. Biol. Chem.* **268**, 25735–25739.

39. Walker, P. A., Leong, L. E. & Porter, A. G. (1995). Sequence and structural determinants of the interaction between the 5′-noncoding region of picornavirus RNA and rhinovirus protease 3C. *J. Biol. Chem.* **270**, 14510–14516.

40. Zhang, X. N., Song, Z. G., Jiang, T., Shi, B. S., Hu, Y. W. & Yuan, Z. H. (2010). Rupintrivir is a promising candidate for treating severe cases of Enterovirus-71 infection. *World J. Gastroenterol.* **16**, 201–209.

41. Kabsch, W. (2010). XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 125–132.

42. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.

43. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.

44. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N. *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–221.