

RESEARCH ARTICLE

An open conformation determined by a structural switch for 2A protease from coxsackievirus A16

Yao Sun^{1,2*}, Xiangxi Wang^{1,2*}, Shuai Yuan^{1,2}, Minghao Dang^{1,2}, Xuemei Li¹, Xuejun C. Zhang¹, Zihao Rao¹✉

¹ National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

² University of Chinese Academy of Sciences, Beijing 100049, China

✉ Correspondence: raozh@xtal.tsinghua.edu.cn

Received August 21, 2013 Accepted September 10, 2013

ABSTRACT

Coxsackievirus A16 belongs to the family *Picornaviridae*, and is a major agent of hand-foot-and-mouth disease that infects mostly children, and to date no vaccines or antiviral therapies are available. 2A protease of enterovirus is a nonstructural protein and possesses both self-cleavage activity and the ability to cleave the eukaryotic translation initiation factor 4G. Here we present the crystal structure of coxsackievirus A16 2A protease, which interestingly forms hexamers in crystal as well as in solution. This structure shows an open conformation, with its active site accessible, ready for substrate binding and cleavage activity. In conjunction with a previously reported “closed” state structure of human rhinovirus 2, we were able to develop a detailed hypothesis for the conformational conversion triggered by two “switcher” residues Glu88 and Tyr89 located within the bl12-cl1 loop. Substrate recognition assays revealed that amino acid residues P1', P2 and P4 are essential for substrate specificity, which was verified by our substrate binding model. In addition, we compared the *in vitro* cleavage efficiency of 2A proteases from coxsackievirus A16 and enterovirus 71 upon the same substrates by fluorescence resonance energy transfer (FRET), and observed higher protease activity of enterovirus 71 compared to that of coxsackievirus A16. In conclusion, our study shows an open conformation of coxsackievirus A16 2A protease and the underlying mechanisms for conformational conversion and substrate specificity. These new insights should facilitate the future rational design of efficient 2A protease inhibitors.

KEYWORDS coxsackievirus A16, 2A protease, crystal structure, switcher

INTRODUCTION

Together with enterovirus 71 (EV71), coxsackievirus A16 (CVA16) is the main pathogen of human hand-foot-and-mouth disease (HFMD). Whilst HFMD rarely occurs in adults, it affects infants and children under five years old, and poses a serious public health threat across the Asia-Pacific region. In most cases, HFMD is a self-limiting disease, with the majority of infected persons recovering spontaneously after symptomatic treatment. Only a small portion of infected children develop neurological complications, which may further progress to potentially fatal cardiopulmonary failure, or even death (Organization, 2011). At present, neither efficacious vaccines nor anti-viral drugs are available against human hand-foot-and-mouth disease.

CVA16 is a non-enveloped virus closely related to EV71 (>90% sequence identity), with both viruses belonging to the human enterovirus A subgroup of the *Enterovirus* genus (Oberste et al., 1999). Recently, the crystal structures of EV71 and CVA16 became available (Wang et al., 2012; Ren et al., 2013). Circulating CVA16 belongs to complex recombinant viruses comprising multiple enterovirus A viruses (Zhao et al., 2011). Enteroviruses contain a positive-strand RNA genome that encodes a single polyprotein, which is processed into mature proteins by viral proteases. Two of them, 2A and 3C, are viral proteases that play an important role in the polyprotein processing. In enteroviruses and rhinoviruses the 2A protease (2A^{pro}) cleaves the site between the C-terminal of capsid protein VP1 and the N-terminal of the protease itself, thus sepa-

*These authors contributed equally to the work.

rating the capsid polyprotein from the nonstructural polyprotein (Toyoda et al., 1986). All the other cleavages are performed by either 3C protease (3C^{pro}), or its precursor 3CD (Ypma-Wong et al., 1988). 2A^{pro} of both enteroviruses and rhinoviruses are multifunctional enzymes, which not only plays an important role in virus polyprotein processing, but also participate in the shutting down of the host protein synthesis through cleavage of the host eukaryotic initiation factor 4G (eIF4G) (Etchison et al., 1982; Gradi et al., 1998; Foeger et al., 2002). Cleavage of eIF4G by viral proteases, therefore, inhibits the cap-dependent translation of cellular mRNAs, thus disabling the host cell protein synthesis machinery (Belsham and Sonenberg, 2000). In previous experiments in yeast system, EV71 2A protease functions as a transcriptional activator, therefore, it is believed that enterovirus 2A^{pro} may also play a role in viral replication and/or pathogenesis (Yang et al., 2010). 2A^{pro} from coxsackievirus B4 (CVB4) was shown to be the principal agent in the etiology of heart disease due to its cleavage of dystrophin (Badorff et al., 1999). Furthermore, poliovirus-derived 2A^{pro} is critical for viral RNA replication as well as hijack the splicing and transcriptional system from the host cell by hydrolyzing native cellular proteins (Li et al., 2001; Castelló et al., 2011). Also, upon infection of the host cells, EV71 2A^{pro} can disrupt immune function-related interferon signaling by reducing the level of interferon receptor-1 (Lu et al., 2012).

Although the structures of 2A^{pro} from human rhinovirus 2 (HRV2) and CVB4 have been determined (Petersen et al., 1999; Baxter et al., 2006) to identify the active site and the oxyanion hole, the exact conformational changes of 2A^{pro} in the process that regulates its substrate specificity and initialize its activity remain to be determined. Due to the inadequate space for substrate binding and a single mutation of a nucleophile residue (C110A) at the active site, it is believed that the structures of 2A^{pro} from HRV2 and CVB4 are in “closed” states. Here, we report the crystal structure of a novel recombinant CVA16 2A^{pro} in hexameric form. Our structure reveals an open substrate binding site, which is controlled by two “switcher” residues, Glu88 and Tyr89, both located within the bII2-cII loop. Comparative analysis with HRV2 2A^{pro} allowed us to propose a detailed, evidence-based hypothesis for the conversion from a “closed” state to an “open” state, which is distinctly different from the conformational plasticity mechanism (Cai et al., 2013). Protease activity assay revealed that amino acid residues at P1', P2 and P4 positions are crucial for substrates cleavage, which is in agreement with our structural data. In conclusion, we believe that our structural and functional studies of CVA16 2A^{pro} will assist in the rational design of efficient peptide inhibitors and antiviral compounds targeting viral 2A^{pro} for blocking HFMD viruses replication.

RESULTS

Preparation and characterization of CVA16 2A^{pro}

Active CVA16 2A^{pro} without any redundant amino acid residues was expressed in *E. coli* (Fig. 1A). CVA16 2A^{pro} was

eluted in two separate peaks from the Hitrap Q column (Fig. 1B). While the 2A^{pro} eluted from the first peak was in dimeric form, as verified by size exclusion chromatography, sedimentation velocity experiments using an analytical ultracentrifuge revealed that 2A^{pro} from the second peak formed hexamers (Fig. 1C). To evaluate hexameric 2A^{pro} cleavage activity, GST-eIF4GII, GST-eIF4GII fusion proteins were expressed and used as substrates, and the protease activity of both wildtype (WT) and mutant 2A^{pro} (C110A) were assayed. SDS-PAGE analysis indicates that whilst the hexameric form of 2A^{pro} can carry out cleaving efficiently, it is not a pre-requisite for cleavage activity, since dimeric 2A^{pro} obtained after disassociation of the hexamer can execute cleavage as well. Moreover, the activity of hexamer 2A^{pro} is slightly inhibited by 10 mmol/L EDTA, due to the fact that depletion of Zn²⁺ partially destabilizes the structure (Fig. 1D and 1E).

Structure determination and overall structure of mature CVA16 2A^{pro}

Crystals of CVA16 2A^{pro} diffracted to 1.8 Å, belonging to the space group R32, and two protein molecules (designated A and B) in an asymmetric unit with a solvent content of 55% (corresponding to a Matthews coefficient $V_M = 2.7 \text{ \AA}^3 \cdot \text{Da}^{-1}$) were found (Matthews, 1968). The crystal structure of CVA16 2A^{pro} was determined by using the molecular replacement method, with the crystal structure of HRV2 2A^{pro} (PDB entry: 2HRV) (Petersen et al., 1999) serving as the search model. The final R_{work} value is 0.17 (R_{free} of 0.20) (Table 1). The first 4 residues of the N-terminus in molecule A could not be traced in the electron density map, while the corresponding region in molecule B was found to be well ordered, reflecting non-identical local environments surrounding the two monomers in the asymmetric unit. The last five residues (146–150), which have been implicated in viral RNA replication (Li et al., 2001), are missing in the maps of both monomers, due to absence of interpretable electron density.

CVA16 2A^{pro} shares sequence and structure similarity with the homologous 2A^{pro} found in HRV2 and CVB4 viruses (Fig. 2A) (Gouet et al., 1999; Larkin et al., 2007). The CVA16 2A^{pro} comprises an N-terminal domain of four-stranded β sheet (bI2, cI, eI2 and fI) and a C-terminal domain of six-stranded β barrel (aII, bII1, cII, dII, eII and fII) (Fig. 2B). The topological graph is shown in Fig. 2C. Within the N-terminal domain, two helical turns α 1 (residues Arg20–Leu22) and α 2 (residues Ser36–Arg38) follow the β strands cI and eI2, respectively, and a short α -helix AI (residues Thr24–Asn30) connects α 1 and eI2. The N-terminal domain and C-terminal domain are linked by a long interdomain loop (residues Thr45–Thr60). In the C-terminal domain, an antiparallel β -hairpin, which is comprised of bII2 and one part of the cII strand, locates next to the six-stranded β barrel, and in addition, makes close contact with residues from the N-terminal domain as well. Three highly conserved residues, His21, Asp39 and Cys110 form the active site of CVA16 2A^{pro} (Fig. 2D) and four conserved residues Cys56,

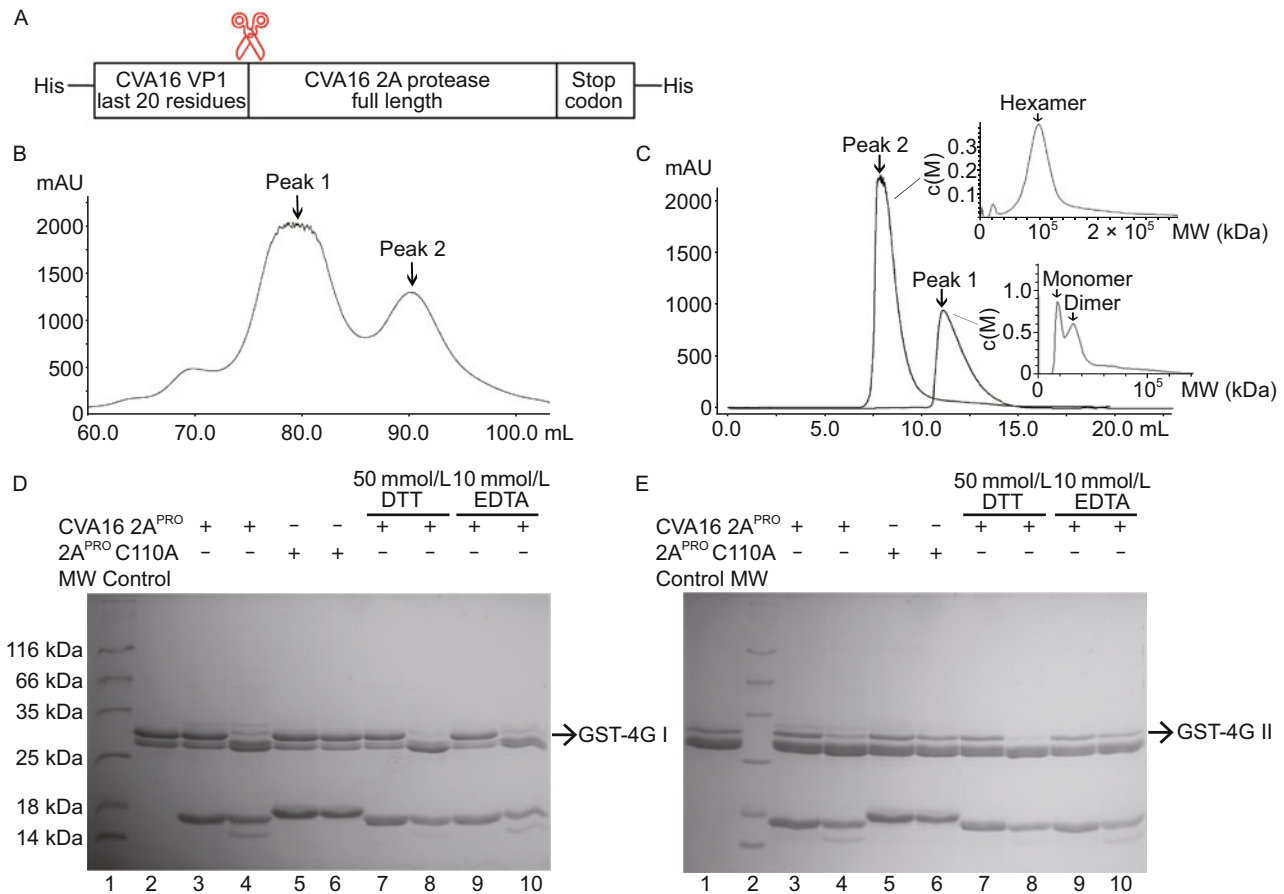


Figure 1. Characterization of recombinant CVA16 2A^{PRO} *in vitro*. (A) The construction strategy of CVA16 2A^{PRO} giving rise to a full length 2A^{PRO} without any redundant amino acids. (B) Two peaks were eluted from the Hitrap Q column and the elution conductances were 34.1 mS/cm and 40.6 mS/cm, respectively. 2 peaks were both the CVA16 2A^{PRO} identified by SDS-PAGE. (C) The size exclusion chromatography and sedimentation velocity experiments showed the two states of CVA16 2A^{PRO}. The first peak was mixture of monomer and dimer while the second peak was hexamer. (D) Cleavage of GST-eIF4GII by CVA16 2A^{PRO} hexamer (lane 3, 4), C110A mutant of CVA16 2A^{PRO} (lane 5, 6), CVA16 2A^{PRO} dimer (lane 7, 8) and CVA16 2A^{PRO} with extra addition of 10 mmol/L EDTA (lane 9, 10) before and after 1 h's reaction. (E) Cleavage of GST-eIF4GII by CVA16 2A^{PRO} hexamer (lane 3, 4), C110A mutant of CVA16 2A^{PRO} (lane 5, 6), CVA16 2A^{PRO} dimer (lane 7, 8) and CVA16 2A^{PRO} with extra addition of 10 mmol/L EDTA (lane 9, 10) before and after 1 h's reaction.

Cys58, Cys116 and His118 make up the zinc ion-binding site (Fig. 2E). Additionally, Cys50 forms an intermolecular disulfide bridge with Cys50 from a symmetry-related molecule.

Comparison with the HRV2 2A^{PRO} and CVB4 2A^{PRO} structures

Exhibiting primary sequence identity of about 36% with HRV2 2A^{PRO} and 73% with CVB4 2A^{PRO}, the overall structure of CVA16 2A^{PRO} chain A shows a high similarity with the HRV2 2A^{PRO} chain A with a root mean square deviation (r.m.s.d) of 1.1 Å for 139 C α atom pairs; and an r.m.s.d of 1.4 Å for 142 C α atom pairs, when compared to CVB4 2A^{PRO}. Unlike the 2 molecules forming an asymmetric unit in the HRV2 2A^{PRO} (with an r.m.s.d of 0.5), the structures of chain A and B are highly similar, with an r.m.s.d of 0.2 Å between 142 equivalent C α atom pairs, and with negligible movements of the loops and

little reorganization of a number of key residue side chains (Fig. 2D). Apart from three extra residues at the N-terminus, and the different orientation of the N- and C-termini in comparison to chain A and B of HRV2 2A^{PRO}, a more unrestricted loop connects bII1 and bII2 β strands in CVA16 2A^{PRO}. In addition, in our structure, cII β strand is one long strand with two functions, one is taking part in the formation of the β barrel, the other is making up the antiparallel β -hairpin together with the bII2. In contrast, in HRV2 the cII β strand consists of two short strands, with one specific function each. The major differences between the two 2A^{PRO} structures occur within the conjunction between α 1 and eI2 of the N-terminal domain (residues Thr24–Asn30 in CVA16 2A^{PRO}) (Fig. 3A–C). When compared to the HRV2 2A^{PRO} sequence, the N-terminal domain of CVA16 2A^{PRO} possesses an α -helix (AI) and an additional amino acid residue (His25) (Fig. 2A). The equivalent regions in the HRV2 A and B chains

Table 1. Data collection and refinement statistics

Name	CVA16 2A ^{PRO} (PDB code: 4MG3)
Data collection	
Resolution (Å)	50.00–1.80 (1.86–1.80)
Space group	R32
Cell dimensions	
a (Å)	101
b (Å)	101
c (Å)	193
α (°)	90.00
β (°)	90.00
γ (°)	120.00
Redundancy	11.1 (10.7)
Completeness (%)	99.8 (100.0)
<i>R</i> _{merge}	0.088 (0.339)
<i>I</i> /σ (<i>I</i>)	20.77 (8.44)
Refinement	
Resolution (Å)	1.80
No. reflections	35, 773
<i>R</i> _{work} / <i>R</i> _{free}	0.17/0.20
No. of non-H atoms	
Protein	2216
Ligand/ion	26
Water	264
B-factors (Å ²)	
Protein	23
Ligand/ion	45
Water	36
Ramachandran statistics (%)	
Most favored	98.9
Allowed	1.1
Outliers	0
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.9

Values in parentheses are for highest-resolution shell.

$$^a R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I(hkl)_i - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I(hkl)_i}$$

$$^b R_{\text{work}} = \frac{\sum_{hkl} |F_o(hkl) - F_c(hkl)|}{\sum_{hkl} F_o(hkl)}$$

^c *R*_{free} was calculated for a test set of reflections (5%) omitted from the refinement.

adopt a 3₁₀ helix and unrestricted loop, respectively. This observation matches the fact that the B factors for this region are high.

The first set of coordinates from the CVB4 2A^{PRO} solution structure (PDB entry: 1Z8R) was selected for subsequent structural analysis in comparison with CVA16 2A^{PRO} (Fig. 3D). Among the differences, the largest conformational shift is the region (residues 103–112) connecting cII and dII, which is believed to be crucial for the activity of protease, with the key amino acid Cys110 located there. CVB4 2A^{PRO} contains an

extended long loop (CD loop) within that region. Unlike the α3 helical turn in CVA16 2A^{PRO}, CDI (cII-dII-1 loop) (residues 102–105) is pulled away from the top of the β barrel by up to 7 Å, which results in CDII (cII-dII-2 loop) (residues 109–112) moving down by 5 Å. The shift of the CDII loop, together with the reorganization of the His21 and Asp39 side chains, results in an increase in distance from 3.8 Å to 11.9 Å between the active sites C110A and His21; in addition, the distance between C110A and Asp39 also extends to 12.7 Å (Fig. 3D). Furthermore, a downward shift of the bII2-cII loop in CVB4 2A^{PRO} by ~5 Å results in a narrowing of the potential substrate binding cleft.

Glu88, Tyr89 as “Switcher” for the substrate binding cleft

There is an open cleft across the surface of CVA16 2A^{PRO}, with sufficient space to accommodate P5-P2' substrate residues, where the active sites are located on the right (Fig. 4A). In HRV2 2A^{PRO}, the cleft becomes narrow, and is filled or partially blocked by the side chains of two residues (Fig. 4B). In contrast, no notable cleft at the substrate-binding site was observed for the CVB4 2A^{PRO} structure (Fig. 4C). Five residues of CVA16 2A^{PRO}, namely Glu85, Glu88, Tyr89, Asn129 and Pro107 (corresponding to Gln81^{HRV2}, Glu84^{HRV2}, Tyr85^{HRV2}, Asp125^{HRV2} and Pro103^{HRV2}), function as gatekeepers to control the size of the cleft. Two of those, Glu88 and Tyr89 (Glu84^{HRV2} and Tyr85^{HRV2}), act as gatekeepers with a “switcher” function: they can convert the cleft conformation from a “closed” to an “open” state, in a reversible manner. Importantly, Glu88, Tyr89 and Tyr90 residues are all highly conserved in picornaviruses. In contrast to CVA16 2A^{PRO}, the side chain of Glu84 in HRV2 2A^{PRO} swings directly towards to cleft. This movement together with a downward shift of the bII2-cII loop by ~1.7 Å, results in both a decrease in distance between gatekeepers Glu and Pro from 7.6 Å in CVA16 2A^{PRO} to 3.5 Å in HRV2 2A^{PRO}, and the formation of an “EP bridge” above the cleft. Meanwhile, the Tyr85 side chain in HRV 2A^{PRO} is rotated clockwise by 90°, in the process of which the cleft becomes partially filled. In addition, the active site residues His18 and Cys106 underneath the rotated Tyr85 side chain are now inaccessible for the substrate P' position (Fig. 4B). In CVA16 2A^{PRO}, in contrast, the cleft presents an open conformation to accommodate substrate, leaving the active site accessible and ready to cleave substrate. In addition, the Asn129/Asp125 residues, located at the south of cleft, possess a variable conformation to regulate the width of the cleft (Fig. 4A and 4B). The fact that the bII2-cII loop residues in both 2A^{PRO}, in particular the two “switcher” (Glu88/Glu84, Tyr89/Tyr85) exhibit higher temperature factors (Fig. 4D), matches the observation that the side chains of two “switcher” are reorganized, and that the bII2-cII loop shifts slightly (Fig. 4A and 4B). On the basis of these findings, we propose a model to describe a mechanism for conformational conversion that Glu88 and Tyr89 function as “switcher”, together with the bII2-cII loop motion, to determine 2A^{PRO} existing as an “open” conformation or a “closed” conformation in the process of initializing its cleavage activity.

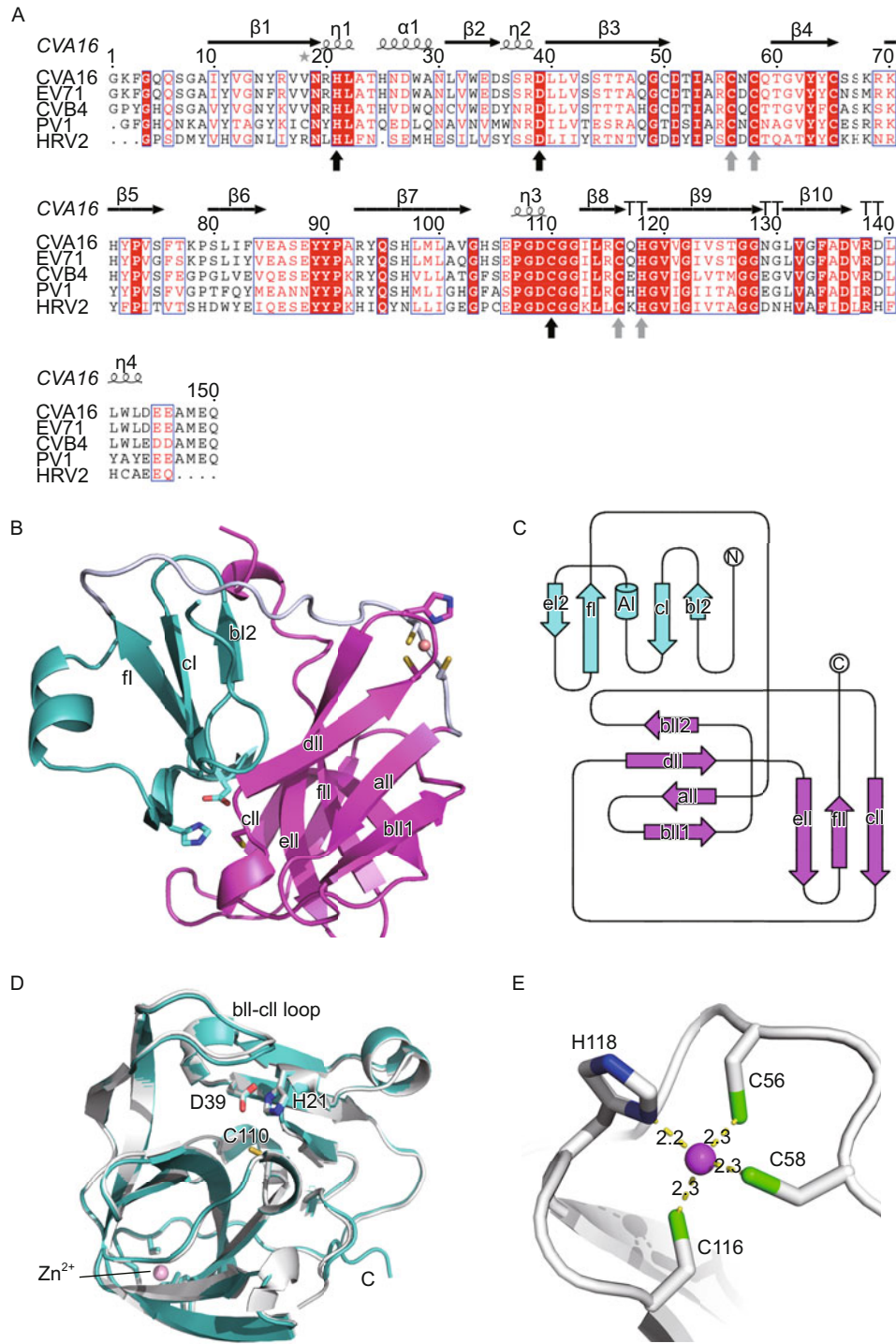


Figure 2. Sequence information and overall structure of CVA16 2A^{PRO}. (A) Amino acid sequences of 2A^{PRO} from coxsackievirus A16, human enterovirus 71, coxsackievirus B4, poliovirus 1 and human rhinovirus 2 were aligned with the program CLUSTALW2 (Larkin et al., 2007), and the alignment result was graphically displayed with the program ESPript (Gouet et al., 1999). Secondary structure elements are indicated according to the structure of CVA16 2A^{PRO}. The black arrows show the conserved active site and the gray arrows show the Zn²⁺ binding site. (B) Overall structure of CVA16 2A^{PRO} chain B. The N-terminal domain and the C-terminal domain of CVA16 2A^{PRO} are shown in cyan and magenta respectively and the interdomain loop is in gray. The zinc ion is represented as a pink sphere. (C) The topological graph of CVA16 2A^{PRO}. (D) Superposition of the CVA16 2A^{PRO} chain A and chain B. The structure of chain A is colored in gray and the chain B is colored in cyan. Active site residues are shown as sticks. (E) A tetrahedral coordination site for the zinc ion binding in CVA16 2A^{PRO} structure.

Protein & Cell

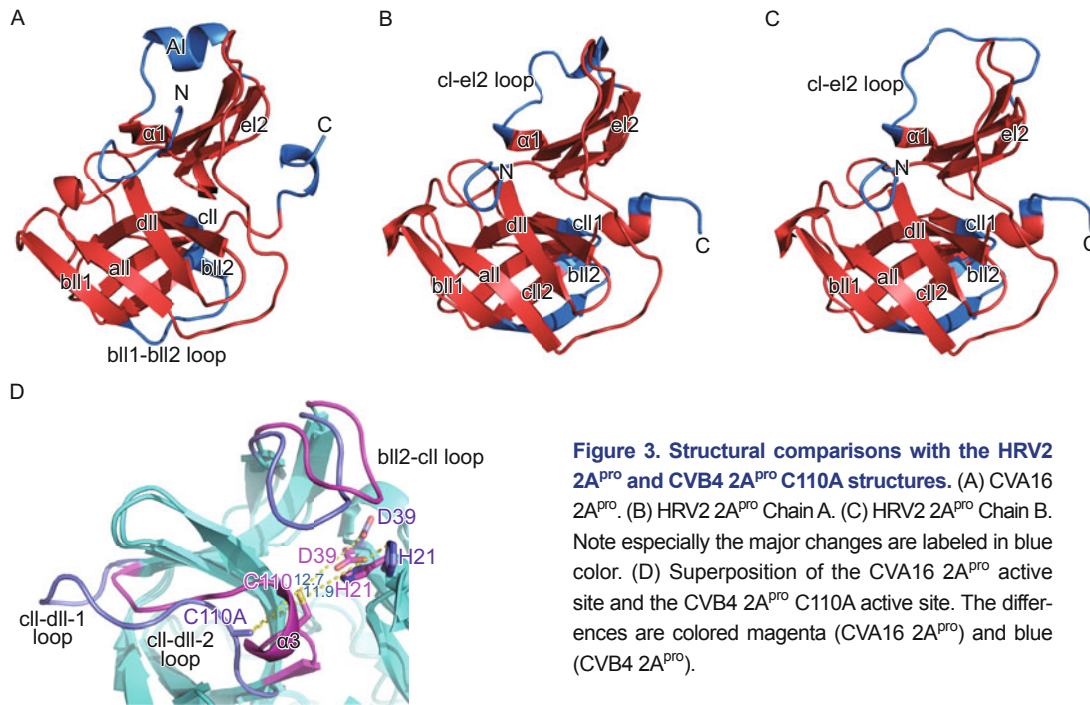


Figure 3. Structural comparisons with the HRV2 2A^{pro} and CVB4 2A^{pro} C110A structures. (A) CVA16 2A^{pro}. (B) HRV2 2A^{pro} Chain A. (C) HRV2 2A^{pro} Chain B. Note especially the major changes are labeled in blue color. (D) Superposition of the CVA16 2A^{pro} active site and the CVB4 2A^{pro} C110A active site. The differences are colored magenta (CVA16 2A^{pro}) and blue (CVB4 2A^{pro}).

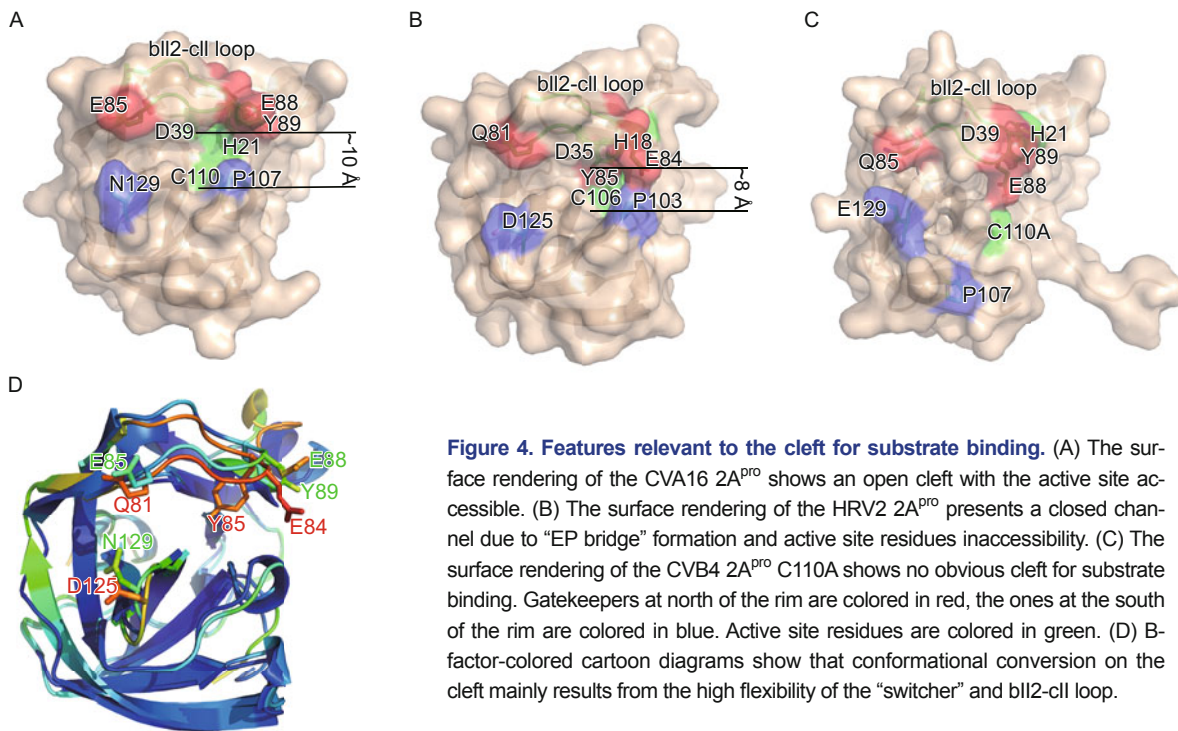


Figure 4. Features relevant to the cleft for substrate binding. (A) The surface rendering of the CVA16 2A^{pro} shows an open cleft with the active site accessible. (B) The surface rendering of the HRV2 2A^{pro} presents a closed channel due to “EP bridge” formation and active site residues inaccessibility. (C) The surface rendering of the CVB4 2A^{pro} C110A shows no obvious cleft for substrate binding. Gatekeepers at north of the rim are colored in red, the ones at the south of the rim are colored in blue. Active site residues are colored in green. (D) B-factor-colored cartoon diagrams show that conformational conversion on the cleft mainly results from the high flexibility of the “switcher” and bll2-cll loop.

Substrate recognition assay

Self-cleavage sites of five types of picornaviral VP1-2A and four identified substrates (or pseudo-substrate) by 2A^{pro} were aligned to determine the amino acids sequence specificity for cleavage (Fig. 5A). The substrate P1’ site, Gly, is the most conserved residue in all substrates aligned, and acts as the actual

cleavage site. The polar amino acid residues Thr/Asn and the nonpolar amino acid residues Leu/Ile are highly conserved at positions P2 and P4, respectively. No significant restrictions were observed on amino acid residues at other positions around the cleavage site.

In order to investigate the importance of the conserved amino acid residues at P1’, P2 and P4 positions for sub-

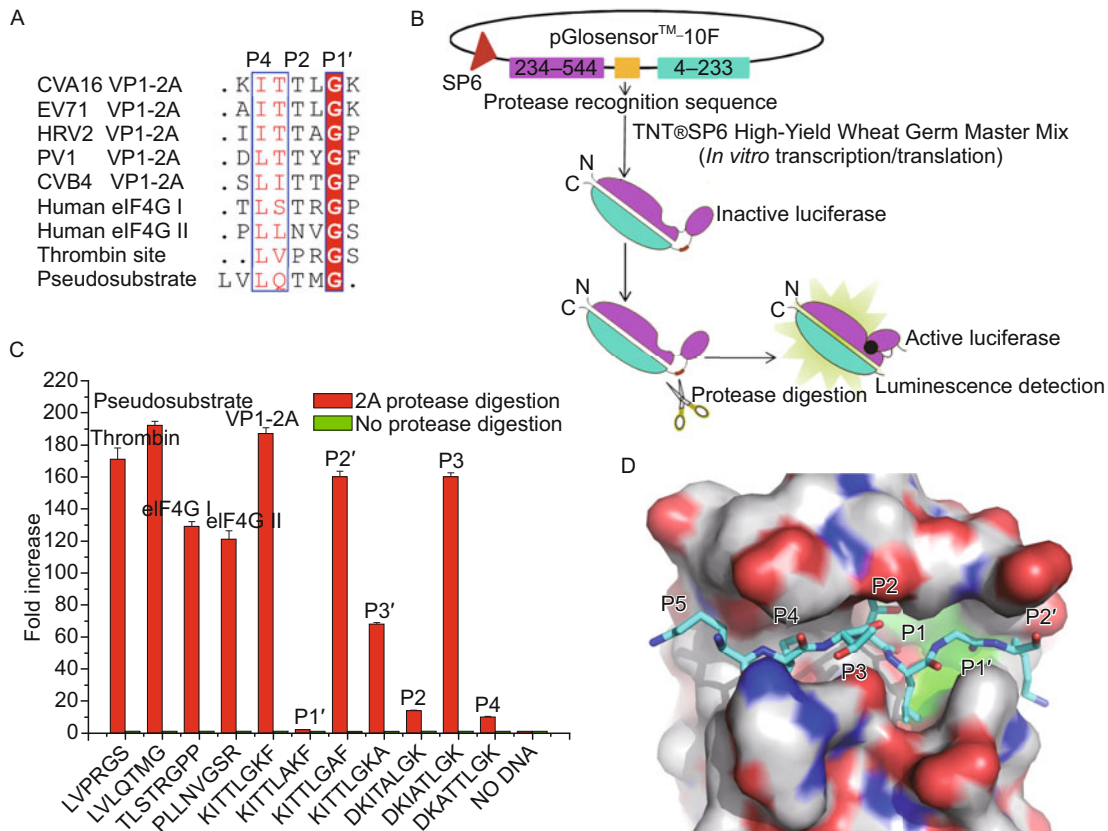


Figure 5. Substrates recognition and substrates binding model. (A) Multiple sequence alignment of several representative 2A^{pro} substrates (or pseudo-substrate). (B) Schematic diagram of the Protease-GloTM Assay of CVA16 2A^{pro}. (C) Protease-GloTM Assay of CVA16 2A^{pro} on *in vitro* translated substrates thrombin site, pseudo-substrate, eIF4GI, eIF4GII, VP1-2A and its mutants. (D) The substrate P5-P2' binding model onto CVA16 2A^{pro}.

strates recognizing and cleavage, a Protease-GloTM Assay was performed. Eleven oligonucleotides encoding different known protease recognition sequences were designed, together with a number of single site mutants, and then cloned into the GloSensorTM-10F gene located on a linearized vector (Fig. 5B). In this assay, if CVA16 2A^{pro} recognizes and cleaves its substrates, which results in separation of the recombinant luciferase and activation of luciferase enzyme, and the luminescence signal increases upon addition of luciferase substrate. Differential cleavage activity with various substrates sequences was detected by luciferase fluorescence, and key amino acid residues crucial to substrates recognition could be identified as well. As shown in Fig. 5C, significant luciferase activity was detected for VP1-2A, eIF4GI, eIF4GII, thrombin recognition site and LVLQTM/G peptide, a potential inhibitor of HRV2 2A^{pro} (Falah et al., 2012). For the self-cleavage substrates of CVA16 2A^{pro}, mutation at P1', P2, P4 positions to Ala decreased CVA16 2A^{pro} cleavage activity dramatically, in particular, the P1' amino acid residue. In contrast, mutations at the positions P2', P3 and P3' have a moderate impact on substrates recognition. Based on these results, we conclude that the highly conserved amino acid residues at the positions P1',

P2 and P4 of the substrate play an essential role in substrate recognition of CVA16 2A^{pro}.

Substrate binding model

To gain further insights into the interaction between CVA16 2A^{pro} and its substrates, and to analyze the specificity for cleavage activity of the amino acid residues at positions P1', P2 and P4, an oligopeptide substrate was modeled virtually into the substrate binding site (Seeliger and de Groot, 2010). The complex structure of CVA16 3C^{pro} and its P1-P10 substrate (PDB entry: 3S9J) was used as a basis for the initial modeling with the sequence KITTL/GK (P5-P2') instead. As shown in Fig. 5D, the results indicate that the P1' residue of Gly locates at the mouth of the channel between the two "switcher" Glu88 and Tyr89, with the active site residues Cys110 and His21 located right underneath. The narrow channel at position P1' presents a steric hindrance to any amino acid residue larger than Gly. The P2 residue, Thr, is placed in the broader pocket. Even though there is enough space to accommodate side chains larger than Thr, the local hydrogen bond network formed by CVA16 2A^{pro} residues Arg93, Gln95 as well as molecular wa-

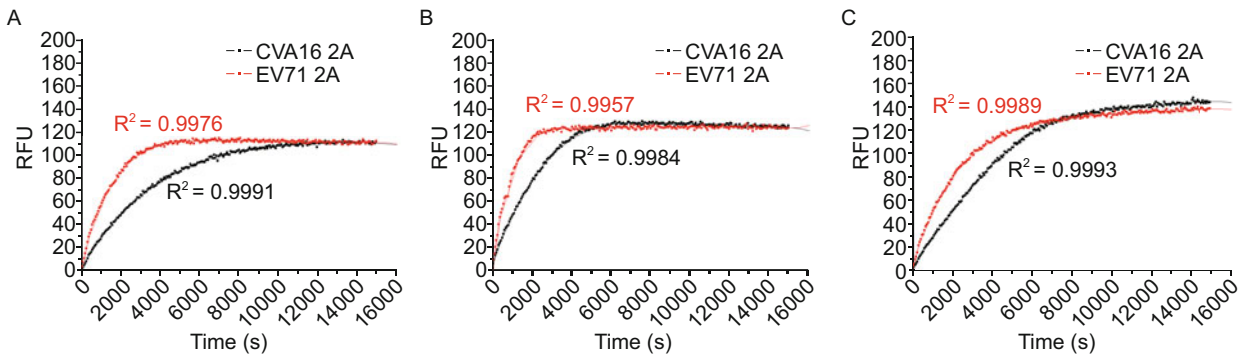


Figure 6. Comparison of CVA16 2A^{pro} activity and EV71 2A^{pro} activity by FRET. (A) Cleavage curve of CVA16 VP1-2A (DabcyL-RDKITTLGKFGQD-EDANS) by CVA16 and EV71 2A^{pro}. The red curve shows the cleavage of VP1-2A peptide by EV71 2A and the black curve shows the cleavage of VP1-2A by CVA16 2A. The fluorescence value is recorded every minute. Origin 7.0 is used to fit the curves. (B) Cleavage curve of human eIF4GI (DabcyL-RTTLSTRGPPRGD-EDANS) by CVA16 and EV71 2A^{pro}. (C) Cleavage curve of human eIF4GII (DabcyL-GVPLLNVGSRRSD-EDANS) by CVA16 and EV71 2A^{pro}.

ter, requires the P2 residue to be a polar amino acid. The side chain of the P4 residue Ile is located inside a deeper cavity, where Ile82, Val84, Leu98 and Leu100, together with a non-polar amino acid residue at P4 position build a hydrophobic pocket. However, the P2', P3 and P4 residues face towards the outside of the cleft, which explains the absence of significant specificity towards amino acid residues for enzymatic activity. Although the P1 residue of Leu is placed inside the cavity, which is enough to accommodate other residues, this could explain in part the lack of specificity at the P1 position.

Comparing the 2A^{pro} activity between the 2 major HFMD viruses by FRET method

In order to test the 2A^{pro} activity of two major HFMD pathogens *in vitro*, we synthesized three modified substrate peptides and performed FRET experiments. As shown in Fig. 6, the reaction of EV71 2A^{pro} could reach the fluorescence maximum faster than CVA16 2A^{pro} in both the reaction of self-cleavage substrate VP1-2A (Fig. 6A) and the reaction of host eIF4GI (Fig. 6B) and eIF4GII (Fig. 6C). The more efficient protease activity for cleaving the VP1-2A substrate of EV71 2A^{pro} than CVA16 2A^{pro} may reflect a faster polyprotein processing, probably leading to a shorter life cycle of EV71 compared to CVA16. Importantly, this increase in cleavage efficiency of eIF4G probably results in higher virulence of EV71, by shutting down protein expression of the host cell. In addition, we find that both 2A proteases can cleave eIF4GI more efficiently than eIF4GII, which is consistent with our results for the Protease-Glo™ Assay (Fig. 5C).

DISCUSSION

Enterovirus proteases 2A^{pro}, together with 3C^{pro}, are responsible for processing polyprotein precursors to produce mature structural proteins and nonstructural viral proteins. On the other hand, enteroviruses 2A^{pro} can shut off host cell protein expression by cleaving eIF4GI, eIF4GII and PABP (Joachims et al.,

1999). Recently, an innate immunity signaling molecule MVAS was identified as a target of EV71 viral protein 2A^{pro}, which cleaved MVAS at multiple residues, ultimately resulting in low expression of type I IFN (Wang et al., 2013).

A monomer is the prominent form for recombinant 2A^{pro}, and the monomer functioning as a dependent active unit has been reported previously (Lamphear et al., 1993; Liebig et al., 1993; Sommergruber et al., 1994; Baxter et al., 2006). Here, we report a novel recombinant CVA16 2A^{pro} that exists in solution as a hexamer, and which disassociates into dimers upon adding of DTT, indicating that disulfide bonds participate in the hexamer formation. Furthermore, in our structural analysis of CVA16 2A^{pro}, we observed intermolecular disulfide bridges formed by Cys50 with its symmetry-related molecule Cys50. The molecular contacts of CVA16 2A^{pro} in the crystal was investigated and analyzed using the PISA program in the CCP4 software package (Krissinel and Henrick, 2007). There are two shapes of hexamers at one plane in the CVA16 2A^{pro} crystal (Fig. 7A), one of which is a standard triangle (Fig. 7B) and the other is a circular shape (Fig. 7C). Taking into account the presence of an intermolecular disulfide bridge, we conclude that the only possible forms that can exist in solution are circular hexamers. PISA analysis indicates that the protein interaction surface area in an asymmetric unit of one molecule A and one molecule B is 516 Å², and the interaction involves three hydrogen bonds and a number of van de Waals interactions. The protein interaction surface area between one molecule A and one symmetric related molecule B is 177.7 Å², and the interaction is established by three hydrogen bonds and one disulfide bond. Similar to the enterovirus protease 2A^{pro} monomer, hexameric CVA16 2A^{pro} also exhibited high cleavage activity. However, one unresolved question concerns the reason for the formation of hexamers by the recombinant CVA16 2A^{pro}. Whilst it is not known whether the hexamer of 2A^{pro} exists under physiological conditions, we can only assume that the 2A^{pro} hexamer possesses specific functions, which cannot

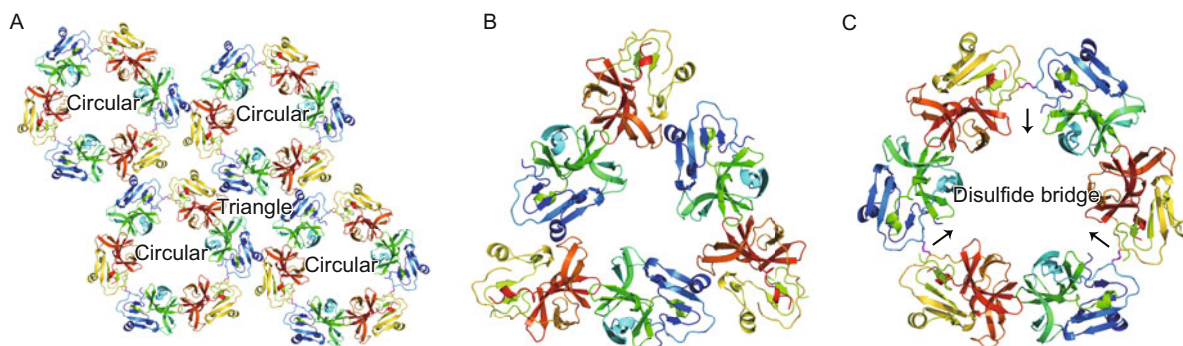


Figure 7. CVA16 2A^{PRO} crystal packing pattern. (A) Two types of hexamers exist in CVA16 2A^{PRO} crystal packing named circular shaped hexamer and triangle shaped hexamer. (B) A standard triangle shaped hexamer without disulfide bridge in hexamer formation. (C) A circular shaped hexamer with disulfide bridge (colored with magenta) in hexamer formation.

be performed by its monomeric or dimeric forms. Further biochemical investigations should shed light on the precise function of the hexamer.

In summary, the CVA16 2A^{PRO} crystal structure presents an open conformation for substrates binding and cleavage, which is mainly controlled by two “switcher”, Glu88 and Tyr89. High temperature factors of the two “switcher” residues, in combination with the bII2-cII loop, render the cleft into a variable conformation for substrate binding. In addition, we analyzed several substrate binding sequences of CVA16 2A^{PRO}, and found the residues at P1', P2, P4 positions to be highly conserved amongst these substrates. Our biochemical data suggest that a single site mutation to Ala at position P1', P2 and P4 decreases the cleavage activity of 2A^{PRO} dramatically, which was verified by our substrate-binding model. Enterovirus proteases 2A^{PRO} are key targets for antiviral drug development, and our study will provide important information for efficient therapeutic invention.

MATERIALS AND METHODS

Protein expression, purification and crystallization

The DNA construct coding for the C-terminal 20 amino acids of CVA16 VP1 fused to 2A^{PRO} sequence was commercially synthesized and inserted into the pET-28a vector between *EcoRI* and *XhoI* restriction sites. Competent *E. Coli* BL21 (DE3) cells were transformed with the recombinant plasmid. The bacterial cultures were grown at 37°C to an OD₆₀₀ of 0.6 and 0.5 mmol/L isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce the 2A^{PRO} at 16°C for 18 h.

Cells were harvested and sonicated in the lysis buffer (20 mmol/L Tris-HCl (pH 8.0), 50 mmol/L NaCl, 10% (v/v) glycerol). The lysate was centrifuged at 30,700 *g* to remove cell debris. The supernatant was added onto a Ni-NTA agarose column (Qiagen) equilibrated with the lysis buffer, and 2A^{PRO} was eluted with the lysis buffer supplemented with 50 mmol/L imidazole. The 2A^{PRO} samples were further purified by ion exchange chromatography with a Hitrap Q column (GE Healthcare). Two peaks were obtained, and both were determined to contain 2A^{PRO} by SDS-PAGE. Fractions of the 2 peaks were collected and concentrated before being loaded onto a Superdex 75 gel filtration column (GE Healthcare) equilibrated with the lysis buffer. The final purified samples

were concentrated individually to ~8 mg/mL.

Crystals of CVA16 2A^{PRO} were grown at 16°C using the hanging-drop vapor diffusion method over a reservoir of 0.1 mol/L MgCl₂·6H₂O, 0.1 mol/L HEPES-Na (pH 7.5) and 18% (w/v) polyethylene glycol (PEG) 400. Crystals were obtained from the 2nd peak of 2A^{PRO} only, and were grown to full size within 3 days.

For mutation, the full length CVA16 2A^{PRO} cDNA sequence was inserted into the pET-28 vector between the restriction sites *NcoI* and *HindIII*. Site-directed mutagenesis of the CV16 2A^{PRO} C110A was performed using the TaKaRa MutanBEST Kit. The expression and purification protocols of the CVA16 2A^{PRO} C110A were identical to that of native protein.

For the protease activity assay, the constructs and expression of the GST-fusion eIF4G fragments were performed as previously reported (Muto et al., 2006). Glutathione Sepharose 4B (GE Healthcare) columns were used to purify the two GST fusion fragments. Proteins were eluted with the lysis buffer containing 15 mmol/L glutathione, and purity of the samples was verified by SDS-PAGE.

The cDNA coding for the EV71 2A^{PRO} was cloned into the pGEX-6p-1 vector with five EV71 VP1 residues at the N-terminal of the 2A^{PRO}. The expression and purification protocols were similar to that of GST-fusion eIF4G fragments.

Analytical ultracentrifugation

Sedimentation velocity experiments were performed on a Beckman XL-I analytical ultracentrifuge at 20°C. Protein samples were diluted with lysis buffer to 400 μL at an A_{280nm} absorption of about 0.75. Samples were loaded into a conventional double-sector quartz cell and mounted in a Beckman four-hole An-60 Ti rotor. Data were collected at 50,000 rpm at a wavelength of 280 nm. Interference sedimentation coefficient distributions were calculated from the sedimentation velocity data using the SEDFIT software program (www.analyticalultracentrifugation.com).

X-ray data collection, processing, and structure determination

Cryo-cooling was carried out by soaking the crystals in reservoir solution containing 20% (v/v) glycerol prior to flash-cooling with liquid nitrogen. All the diffraction data sets were collected at beam line BL1A of the Photon Factory (PF) synchrotron facility in Japan using a CCD

detector, with highest resolution at ~ 1.8 Å. The space group of CVA16 2A^{pro} crystals is R32, with 2 protein molecules per asymmetric unit (ASU). The initial phases were obtained with the molecular replacement method by the program Phaser v2.1 (McCoy et al., 2007), and the HRV2 2A^{pro} crystal structure was used as the search model (Protein Data Bank (PDB) ID: 2HRV). Manual model building and refinement were performed with COOT (Emsley and Cowtan, 2004) and PHENIX (Adams et al., 2010) following rigid body, TLS parameter and individual B-factor refinement. Solvent molecules were located from stereochemically reasonable peaks in the σ_A -weighted 2Fo-Fc difference electron density map. The quality of the final refined model was verified using the program PROCHECK (Laskowski et al., 1993). Final refinement statistics are summarized in Table 1. Structural figures were drawn with the program PyMOL (DeLano, 2002).

Protease activity detection by Protease-Glo™ Assay

Cleavage activity of the recombinant proteases on a number of substrate peptides and mutant peptides were evaluated *in vitro* using a Protease-Glo™ Assay (Promega). This assay detects and measures protease activities using a genetically engineered firefly (*Photinus pyralis*) luciferase. Eleven oligonucleotides encoding protease recognition sequences or mutants were designed and cloned into the GloSensor™-10F gene located on a linearized vector. The peptides used are LVPRGS (the thrombin site), LVLQTMG (the reported pseudo-substrate of HRV2 2A^{pro}), TLSTRGPP (eIF4GI recognition site), PLLNVGSR (eIF4GII recognition site), KITTLGKF (VP1-2A self-cleavage site), KITTLAKF (P1' mutant), KITTLGAF (P2' mutant), KITTLGKA (P3' mutant), DKITALGK (P2 mutant), DKIATLGK (P3 mutant) and DKATTLGK (P4 mutant). After *in vitro* proteins synthesis for 2 h, tested protease was added into the reaction solution and incubated for 1 h at 30°C. Luminescence was measured using a luminometer 5 min after addition of Bright-Glo assay reagent.

The synthesis of the chemically modified peptides and the FRET experiment for activity comparison

The peptides used for FRET experiments were synthesized at Sci-Light Biotechnology, LIC. The donor group EDANS was chemically linked to the carboxyl terminal of the peptides having an extra Asp residue. The acceptor group Dabcyl was linked to the amino terminal of the peptides. Three peptides were synthesized as follows: Dabcyl-RTTLSTRGPP RGD (EDANS)-NH₂ (the eIF4GI peptide), Dabcyl-GVPLLNVGSR RSD (EDANS)-NH₂ (the eIF4GII peptide) and Dabcyl-RDKITTLGKF GQD (EDANS)-NH₂ (the VP1-2A peptide). The peptides were dissolved in dimethylsulfoxide (DMSO) to a concentration of 10 mg/mL and were stored at -20°C prior to use.

All three peptides were diluted to 100 mmol/L, and the molar ratio of the tested protease to peptide was 1:20. After adding the protease to the substrate peptide, the fluorescence values were read every minute, using a 96-well black plate, with 3 replicates per reaction.

ACCESSION CODES

The coordinate and structure factor are deposited with the RCSB under accession codes: 4MG3.

ACKNOWLEDGEMENTS

We gratefully acknowledge the assistance of the staff of the beamline

BL1A at the Photon Factory (PF) in Japan with the X-ray diffraction data collection. This work was supported by grants from the National Basic Research Program (973 Program) (Nos. 2014CB542800 and 2011CB915501), the National Natural Science Foundation of China (Grant No. 31170702). We thank Ms. Ya Wang for technical support with the FRET experiments, and Ms. Xiaoxia Yu for her excellent assistance with the analytical ultracentrifugation from the Structural Biology Core Facility in the Institute of Biophysics, Chinese Academy of Sciences (CAS).

ABBREVIATIONS

ASU, asymmetric unit; CVA16, coxsackievirus A16; CVB4, coxsackievirus B4; DMSO, dimethylsulfoxide; DTT, dithiothreitol; *E. Coli*, *Escherichia coli*; EDTA, ethylene diamine tetraacetic acid; eIF4G, eukaryotic initiation factor 4G; EV71, enterovirus 71; FRET, fluorescence resonance energy transfer; GST, glutathione S-transferase; HFMD, hand-foot-and-mouth disease; HRV2, human rhinovirus 2; IPTG, isopropyl β -D-1-thiogalactopyranoside; LB, Luria-Bertani; PDB, Protein Data Bank; PEG, polyethylene glycol; *r.m.s.d.*, root mean square deviation; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis

COMPLIANCE WITH ETHICS GUIDELINES

Yao Sun, Xiangxi Wang, Shuai Yuan, Minghao Dang, Xuemei Li, Xuejun C. Zhang and Zihao Rao declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by the any of the authors.

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66, 213–221.
- Badorff, C., Lee, G.H., Lamphear, B.J., Martone, M.E., Campbell, K.P., Rhoads, R.E., and Knowlton, K.U. (1999). Enteroviral protease 2A cleaves dystrophin: evidence of cytoskeletal disruption in an acquired cardiomyopathy. *Nat Med* 5, 320–326.
- Baxter, N.J., Roetzer, A., Liebig, H.D., Sedelnikova, S.E., Hounslow, A.M., Skern, T., and Waltho, J.P. (2006). Structure and dynamics of coxsackievirus B4 2A proteinase, an enzyme involved in the etiology of heart disease. *J Virol* 80, 1451–1462.
- Belsham, G.J., and Sonenberg, N. (2000). Picornavirus RNA translation: roles for cellular proteins. *Trends Microbiol* 8, 330–335.
- Cai, Q., Yameen, M., Liu, W., Gao, Z., Li, Y., Peng, X., Cai, Y., Wu, C., Zheng, Q., and Li, J. (2013). Conformational plasticity of the 2A proteinase from enterovirus 71. *J Virol* 87, 7348–7356.
- Castelló, A., Álvarez, E., and Carrasco, L. (2011). The multifaceted poliovirus 2A protease: regulation of gene expression by picornavirus proteases. *J Biomed Biotechnol* 2011:369648.
- DeLano, W.L. (2002). The PyMOL molecular graphics system.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126–2132.
- Etchison, D., Milburn, S.C., Ederly, I., Sonenberg, N., and Hershey, J.W. (1982). Inhibition of HeLa cell protein synthesis following po-

- liovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. *J Biol Chem* 257, 14806–14810.
- Falah, N., Violot, S., Decimo, D., Berri, F., Foucault-Grunenwald, M.L., Ohlmann, T., Schuffenecker, I., Morfin, F., Lina, B., Riteau, B., et al. (2012). Ex vivo and in vivo inhibition of human rhinovirus replication by a new pseudosubstrate of viral 2A protease. *J Virol* 86, 691–704.
- Foeger, N., Glaser, W., and Skern, T. (2002). Recognition of eukaryotic initiation factor 4G isoforms by picornaviral proteinases. *J Biol Chem* 277, 44300–44309.
- Gouet, P., Courcelle, E., and Stuart, D.I. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305–308.
- Gradi, A., Svitkin, Y.V., Imataka, H., and Sonenberg, N. (1998). Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *Proc Natl Acad Sci U S A* 95, 11089–11094.
- Joachims, M., Van Breugel, P.C., and Lloyd, R.E. (1999). Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J Virol* 73, 718–727.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372, 774–797.
- Lamphear, B., Yan, R., Yang, F., Waters, D., Liebig, H., Klump, H., Kuechler, E., Skern, T., and Rhoads, R. (1993). Mapping the cleavage site in protein synthesis initiation factor eIF-4 gamma of the 2A proteases from human Coxsackievirus and rhinovirus. *J Biol Chem* 268, 19200–19203.
- Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, H., Valentin, F., Wallace, I., Wilm, A., and Lopez, R. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26, 283–291.
- Li, X., Lu, H.H., Mueller, S., and Wimmer, E. (2001). The C-terminal residues of poliovirus proteinase 2A(pro) are critical for viral RNA replication but not for cis- or trans-proteolytic cleavage. *J Gen Virol* 82, 397–408.
- Liebig, H.D., Ziegler, E., Yan, R., Hartmuth, K., Klump, H., Kowalski, H., Blaas, D., Sommergruber, W., Frasel, L., Lamphear, B., et al. (1993). Purification of two picornaviral 2A proteinases: interaction with eIF-4 gamma and influence on in vitro translation. *Biochemistry* 32, 7581–7588.
- Lu, J., Yi, L., Zhao, J., Yu, J., Chen, Y., Lin, M.C., Kung, H.F., and He, M.L. (2012). Enterovirus 71 disrupts interferon signaling by reducing the level of interferon receptor 1. *J Virol* 86, 3767–3776.
- Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 33, 491–497.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Stornik, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658–674.
- Muto, S., Miyoshi, H., Nishikawa, H., and Nakashima, H. (2006). Novel recognition sequence of coxsackievirus 2A proteinase. *Biochem Biophys Res Commun* 348, 1436–1442.
- Oberste, M.S., Maher, K., Kilpatrick, D.R., and Pallansch, M.A. (1999). Molecular evolution of the human enteroviruses: Correlation of serotype with VP1 sequence and application to picornavirus classification. *J Virol* 73, 1941–1948.
- Organization, W.H. (2011). A guide to clinical management and public health response for hand, foot and mouth disease (HFMD). Geneva: WHO.
- Petersen, J.F., Cherney, M.M., Liebig, H.D., Skern, T., Kuechler, E., and James, M.N. (1999). The structure of the 2A proteinase from a common cold virus: a proteinase responsible for the shut-off of host-cell protein synthesis. *EMBO J* 18, 5463–5475.
- Ren, J., Wang, X., Hu, Z., Gao, Q., Sun, Y., Li, X., Porta, C., Walter, T.S., Gilbert, R.J., and Zhao, Y. (2013). Picornavirus uncoating intermediate captured in atomic detail. *Nat Commun* 4:1929.
- Seeliger, D., and de Groot, B.L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des* 24, 417–422.
- Sommergruber, W., Ahorn, H., Klump, H., Seipelt, J., Zoepfel, A., Fessl, F., Krystek, E., Blaas, D., Kuechler, E., Liebig, H.D., et al. (1994). 2A proteinases of coxsackie- and rhinovirus cleave peptides derived from eIF-4 gamma via a common recognition motif. *Virology* 198, 741–745.
- Toyoda, H., Nicklin, M.J., Murray, M.G., Anderson, C.W., Dunn, J.J., Studier, F.W., and Wimmer, E. (1986). A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. *Cell* 45, 761–770.
- Wang, B., Xi, X., Lei, X., Zhang, X., Cui, S., Wang, J., Jin, Q., and Zhao, Z. (2013). Enterovirus 71 protease 2Apro targets MAVS to inhibit anti-viral type I interferon responses. *PLoS Pathog* 9, e1003231.
- Wang, X., Peng, W., Ren, J., Hu, Z., Xu, J., Lou, Z., Li, X., Yin, W., Shen, X., and Porta, C. (2012). A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. *Nat Struct Mol Biol* 19, 424–429.
- Yang, C.H., Li, H.C., Jiang, J.G., Hsu, C.F., Wang, Y.J., Lai, M.J., Juang, Y.L., and Lo, S.Y. (2010). Enterovirus type 71 2A protease functions as a transcriptional activator in yeast. *J Biomed Sci* 17, 65.
- Ypma-Wong, M.F., Dewalt, P.G., Johnson, V.H., Lamb, J.G., and Semler, B.L. (1988). Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. *Virology* 166, 265–270.
- Zhao, K., Han, X., Wang, G.J., Hu, W., Zhang, W.Y., and Yu, X.F. (2011). Circulating coxsackievirus A16 identified as recombinant type A human enterovirus, China. *Emerg Infect Dis* 17, 1537–1540.