The structure of the 2A proteinase from a common cold virus: a proteinase responsible for the shut-off of host-cell protein synthesis

Jens F.W.Petersen, Maia M.Cherney,
Hans-Dieter Liebig¹, Tim Skern¹,
Ernst Kuechler¹ and Michael N.G.James²

MRC Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, AB, Canada T6G 2H7 and ¹Institute of Biochemistry, Medical Faculty, University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria

²Corresponding author
e-mail: michael.james@ualberta.ca

The crystal structure of the 2A proteinase from human rhinovirus serotype 2 (HRV2-2A^{pro}) has been solved to 1.95 Å resolution. The structure has an unusual, although chymotrypsin-related, fold comprising a unique four-stranded β sheet as the N-terminal domain and a six-stranded β barrel as the C-terminal domain. A tightly bound zinc ion, essential for the stability of HRV2-2A^{pro}, is tetrahedrally coordinated by three cysteine sulfurs and one histidine nitrogen. The active site consists of a catalytic triad formed by His18, Asp35 and Cys106. Asp35 is additionally involved in an extensive hydrogen-bonding network. Modelling studies reveal a substrate-induced fit that explains the specificity of the subsites S4, S2, S1 and S1'. The structure of HRV2-2A^{pro} suggests the mechanism of the *cis* cleavage and its release from the polyprotein.
Keywords: cysteine proteinase/human rhinovirus/serine proteinase/translational control/zinc binding

Introduction

The successful replication of many viruses depends upon proteolytic processing. Most picornaviruses, a family including poliovirus, human rhinovirus and foot-and-mouth-disease virus, encode at least two proteinases that cleave the primary translation product of protein synthesis, generated from the unique single open reading frame of the positive-strand RNA genome (Ryan and Flint, 1997). The processing pattern varies subtly amongst family members. In entero- and rhinoviruses, the initial cleavage is accomplished by the 2A proteinase (2A^{pro}) which cleaves between its own N-terminus and the C-terminus of VP1, thereby separating the capsid protein precursor from that of the non-structural proteins (Toyoda *et al.*, 1986). All of the other processing steps, except for a maturation cleavage occurring during capsid assembly, are carried out by the 3C proteinase (3C^{pro}) and/or its immediate precursor 3CD^{pro} (Dougherty and Semler, 1993).

Like many viral proteins, both 2A^{pro} and 3C^{pro} are multifunctional enzymes. In addition to processing the viral polyprotein, 2A^{pro} may play a role in RNA replication (Molla *et al.*, 1993; Yu *et al.*, 1995). Furthermore, 2A^{pro} specifically cleaves both homologues of the host protein

eukaryotic initiation factor eIF4G (Lloyd *et al.*, 1988; Gradi *et al.*, 1998). The eIF4G homologues are part of the initiation factor complex eIF4F, which also comprises the RNA unwinding protein eIF4A and the cap binding protein eIF4E. This complex is required for recruitment of capped cellular mRNA to the ribosome; cleavage of the eIF4G homologues impairs this process and leads to the so-called host-cell shut-off, i.e. the inability of the cell to initiate protein synthesis on its own mRNA. However, as the ribosomes initiate internally on a higher ordered structure in the 5' UTR of the picornaviral mRNA known as the IRES (internal ribosome entry segment; Jackson *et al.*, 1990; Jang *et al.*, 1990; Belsham and Sonenberg, 1996), picornaviral protein synthesis is not affected. Indeed, IRES-dependent initiation is actually stimulated by 2A^{pro} cleavage of eIF4G (Hambidge and Sarnow, 1992; Liebig *et al.*, 1993; Ziegler *et al.*, 1995; Borman *et al.*, 1997).

Both the 2A^{pro} of rhino- and enteroviruses, and the 3C^{pro} present in all picornaviruses, were predicted to possess a chymotrypsin-like fold by alignment to known structures, even though cysteine is their active site nucleophile (Bazan and Fletterick, 1988; Gorbalenya *et al.*, 1989). Determination of the three-dimensional structures of 3C^{pro} from hepatitis A virus (HAV), human rhinovirus serotype 14 (HRV14) and poliovirus 1 (PV1) confirmed these predictions (Allaire *et al.*, 1994; Matthews *et al.*, 1994; Bergmann *et al.*, 1997; Mosimann *et al.*, 1997). We have determined, by X-ray crystallography, the three-dimensional structure of the wild-type HRV2-2A^{pro} at 1.95 Å resolution. The structure illuminates the proteolytic mechanism of both intramolecular and intermolecular cleavage, shows how HRV2-2A^{pro} coordinates the bound zinc ion and provides a clear explanation for its restricted substrate specificity. Most importantly, however, given the high degree of conservation amongst the 2A^{pro} of rhinoviruses and of enteroviruses, this structure has potential for the design of antiviral drugs, especially as an enteroviral 2A^{pro} from coxsackievirus serotype B3 has recently been shown to cleave heart muscle dystrophin specifically and may thus be implicated in events leading to virally induced myocarditis (Badorff *et al.*, 1999).

Results and discussion

The structure of recombinant mature HRV2-2A^{pro}

Figure 1A shows the overall structure of HRV2-2A^{pro} as determined using the multiwavelength anomalous dispersion (MAD) technique (see Materials and methods and Table I for crystallographic parameters). The asymmetric unit contains two HRV2-2A^{pro} molecules (designated A and B). Each molecule has a structure comprising an N-terminal domain and a C-terminal domain. An assignment of the β strands in HRV2-2A^{pro} is shown in Table II.

Fig. 1. The structure of HRV2-2A^{pro}. (A) Ribbon diagram of the overall structure of HRV2-2A^{pro}. The strands are labelled according to the topology of SGPB (Read *et al.*, 1983). β strands of the N-terminal domain (light blue, Roman numeral I) and the C-terminal domain (dark blue, Roman numeral II) are indicated. Members of the catalytic triad, the zinc ion (purple sphere) and the zinc binding site are shown. (B) Superposition of the A molecule (red) and B molecule (blue) of HRV2-2A^{pro} in the asymmetric unit. Figure 1 as well as Figures 2, 3, 5, 6 and 7 were generated using the program MOLSCRIPT (Kraulis, 1991) and rendered using Raster3D (Merrit and Bacon, 1997), except where otherwise stated.

The N-terminal domain is constructed from four β strands (bI2, cI, eI2 and fI) that form a β sheet. In addition, the first four amino acids form a type I β turn that is well ordered in both molecules. Of the two active site residues found in the N-terminal domain, His18 is situated in a

short 3_{10} helix following strand cI, whereas Asp35 is found in a loop connecting strand eI2 to strand fI. The N-terminal domain is connected to the C-terminal domain by residues 40–52, forming an interdomain loop. A topologically similar loop found in the picornaviral 3C proteinases holds the consensus sequence Lys-Phe-Arg-Asp-Ile, which has been implicated in RNA recognition (Matthews *et al.*, 1994; Bergmann *et al.*, 1997; Mosimann *et al.*, 1997). However, this motif is not found in the interdomain loop of 2A^{pro}. The C-terminal domain itself is a six-stranded β barrel comprising residues 52–132 and is built up of strands aII, bII1, bII2, cII, dII, eII and fII. The nucleophilic Cys106 lies between two consecutive type II β turns (Glu102–Asp105 and Asp105–Gly108) preceding strand dII. A one-turn 3_{10} helix is present at the C terminus (residues 134–137). In addition, Cys138 forms an intermolecular disulfide bridge with Cys138 from a symmetry-related molecule. The last three residues (140–142) could not be traced in the electron density map for either molecule A or B. These are most likely disordered. In enteroviruses, the C terminus has been implicated in viral RNA replication (Lu *et al.*, 1995). Such a correlation has not been described for HRV2-2A^{pro}, which contains 3–4 fewer amino acids.

The two molecules in the asymmetric unit are shown superimposed upon each other in Figure 1B. The root mean square deviation (r.m.s.d.) between 130 equivalent C $_{\alpha}$ atoms is 0.53 Å (cut-off 1.5 Å). The largest conformational difference between the molecules is in the loop from residues 19 to 29, following His18. Surprisingly, this is not the region of the molecule with the highest temperature factors; indeed, these residues in both molecules are reasonably well defined in the electron density. Intermolecular interactions between the β strand eI2 (residues 30–35) in molecule A and an equivalent strand from a symmetry-related B molecule, which follows this region, may explain the differences. These large differences indicate that this loop is flexible in solution. The second largest difference between the A and B molecules is found in a β hairpin loop formed by residues 80–90. This region exhibits the largest temperature factors in both molecules and also has the least well-defined electron density.

Comparison to other chymotrypsin-related proteinases

The fold of HRV2-2A^{pro} has been proposed to be closely related to that of the chymotrypsin-related proteinases (Bazan and Fletterick, 1988; Gorbalenya *et al.*, 1989; Sommergruber *et al.*, 1997). In Figure 2, a structural alignment between molecule A of HRV2-2A^{pro} and the *Streptomyces griseus* proteinase B (SGPB) is shown (Read *et al.*, 1983; PDB accession code 3SGB). SGPB was chosen as a representative structure of a minimal chymotrypsin-like proteinase fold. Although SGPB shares <20% sequence identity with HRV2-2A^{pro}, Figure 2 shows that the two structures align reasonably well with an r.m.s.d. between 50 C $_{\alpha}$ atom pairs of 0.9 Å (cut-off 1.5 Å). The most dramatic differences are found in the N-terminal domain in which only four β strands (bI2, cI, eI2 and fI) superimpose; β strands aI, bI1, dI and eI1 present in the bacterial enzyme are missing in HRV2-2A^{pro} (Figure 2). Thus, one can consider the N-terminal domain of HRV2-2A^{pro} to be a severely truncated β barrel in which

Table I. Data collection, phasing and refinement parameters

Crystal information				
Space group	P3 ₂ 21			
Unit cell dimensions (Å)	<i>a</i> = <i>b</i> = 72.92 and <i>c</i> = 161.66			
Estimated solvent content ^a (%)	69			
Data collection (numbers in parentheses refer to the outermost resolution bin)				
Crystal	1	1	1	2
Beamline	X12C	X12C	X12C	X25
Wavelength (Å)	1.2835 (L1)	1.2827 (L2)	1.2450 (L3)	1.1
Resolution (Å)	2.2	2.2	2.2	1.95
High-resolution range (Å)	2.28–2.20	2.28–2.20	2.28–2.20	2.02–1.95
<i>I</i> / <i>ΣI</i>	28.5	28.3	29.9	15.6
Reflections	140 409	140 810	144 007	109 005
Unique	24 269	24 284	24 879	36 626
Unique anomalous	45 457	45 478	46 658	ND
Completeness (%)	92.6 (71.6)	92.6 (71.6)	94.6 (83.1)	98.5 (96.6)
<i>R</i> _{sym} ^b (%)	5.9 (17.8)	6.4 (18.6)	6.1 (16.2)	5.3
<i>R</i> _{iso} ^c (%)	ND	2.2	4.0	6.9
Phasing statistics to 2.2 Å (acentric/centric)				
<i>R</i> _{cullis} ^d (dispersive)	ND	0.93/0.88	0.85/0.69	not used
<i>R</i> _{cullis} ^d (anomalous)	0.83	0.76	0.81	not used
Phasing power	ND	0.72/0.53	1.23/1.01	not used
Figure of merit	0.44			
Refinement [numbers in parentheses refer to the outermost resolution bin (2.02–1.95 Å)]				
Resolution range (Å)	50–1.95			
<i>R</i> _{work} ^e	0.218 (0.271)			
<i>R</i> _{free} ^e	0.250 (0.299)			
Number of atoms or molecules				
protein	2226			
water	277			
zinc	2			
R.m.s. deviation from ideal geometry				
bond lengths (Å)	0.009			
bond angles (°)	1.5			
Average <i>B</i> (Å ²)	33.15			

^aSolvent content estimated according to Matthews (1968).

$$^b R_{\text{sym}} = \frac{\sum_{hkl} \sum_i |I(h,k,l,i) - \langle I(h,k,l) \rangle|}{\sum_{hkl} \sum_i I(h,k,l,i)}$$

$$^c R_{\text{iso}} = \frac{\sum_{hkl} \|F_2(h,k,l) - F_1(h,k,l)\|}{\sum_{hkl} |F_1(h,k,l)|}$$

$$^d R_{\text{cullis}} = \frac{\sum_{hkl} \|F_{\text{ph}} \pm F_{\text{nat}} - |F_{\text{h(calc)}}|\|}{\sum_{hkl} |F_{\text{ph}} \pm F_{\text{p}}|}$$

$$^e R = \frac{\sum_{hkl} \|F_{\text{obs}}(h,k,l) - k|F_{\text{calc}}(h,k,l)|\|}{\sum_{hkl} |F_{\text{obs}}(h,k,l)|}$$

R_{work} is calculated from a set of reflections where 10% of the total reflections have been randomly omitted from the refinement and used to calculate R_{free} .

only one of the four-stranded β sheets is conserved. The C-terminal domains of the two proteinases are topologically identical. Interestingly, the catalytic triads of the two proteinases also align reasonably well. In general, it seems

that, amongst the chymotrypsin-related proteinases, the N-terminal domain allows for larger topological variations than does the C-terminal domain. In the Sindbis core protein, the six-stranded β barrel of the N-terminal domain

Fig. 2. Comparison of HRV2-2A^{pro} with SGPB. Stereo drawing showing the alignment between SGPB (red) and HRV2-2A^{pro} (blue). β strands are labelled as in Figure 1.

Table II. Assignment^a of β strands in HRV2-2A

Residues	Label
7–8	bI2
8–16	cI
29–31	eI2
36–46	fI
56–61	aII
66–71	bII1
73–80	bII2
89–98	cII
109–112	dII
115–123	eII
127–132	fII

^aThe β strands are defined as in the program PROMOTIF (Hutchinson and Thornton, 1996).

is reduced to a five-stranded β barrel (Choi *et al.*, 1991). Moreover, the hepatitis C virus NS3 proteinase has a peptide cofactor inserted into its N-terminal domain in the place of one of the β strands (Kim *et al.*, 1996; Love *et al.*, 1996). However, HRV2-2A^{pro} is unique as it is the first known chymotrypsin-like proteinase in which the N-terminal domain is not a β barrel, but a four-stranded antiparallel β sheet.

The HRV2-2A^{pro} zinc binding site

Another unusual feature of the picornaviral 2A^{pro} is the presence of a tightly bound zinc ion. The zinc binding site in HRV2-2A^{pro} is found at the beginning of the C-terminal domain. The coordination is tetrahedral; the ligands are three cysteine sulfurs and one histidine nitrogen (Figure 3). Two of the ligands (Cys52S ^{γ} and Cys54S ^{γ}) are donated from the end of the interdomain loop connecting the N- and C-terminal domains, with the two other ligands (Cys112S ^{γ} and His114N ^{δ 1}) lying on a hairpin loop between β strands dII and eII. The four ligands are highly conserved among the known 2A^{pro} sequences (Figure 4). The average ligand to zinc distances are 2.3 Å for the S–Zn bonds and 2.0 Å for the N–Zn bond, similar to those of zinc binding sites in high-resolution structures of similar coordination

Fig. 3. The zinc binding site in HRV2-2A^{pro}. The view is rotated 180° compared with that in Figure 1A. N- (light blue) and C-terminal (dark blue) β strands are shown. Labelled residues are those coordinating the zinc ion (purple sphere). The β strands adjacent to the coordinating residues are labelled.

(Alberts *et al.*, 1998). Studies of the binding of the zinc ion suggest that it is tightly bound and essential for the structure (Sommergruber *et al.*, 1994b; Voss *et al.*, 1995). The zinc ion is not accessible to other ligands; therefore, it is very unlikely that it has any functional role such as binding to RNA. In addition, the distance to the active

Fig. 4. Multiple sequence alignment of picornaviral 2A proteinases. All sequences were retrieved from the Swissprot database. Blue columns, identical residues; red columns, highly conserved residues. hrv, human rhinovirus; pol, poliovirus; svdvu, swine vesicular disease virus; bobev, bovine enterovirus; cox, coxsackievirus; ec, echovirus; he, human enterovirus. Numbers following these codes indicate the serotype.

site is >20 Å, thus making a direct influence of the zinc ion on the proteolytic activity of HRV2-2A^{pro} improbable.

At present, the only other known example of a zinc binding site in a chymotrypsin-related proteinase is that in the NS3 serine proteinase of hepatitis C virus (Kim *et al.*, 1996; Love *et al.*, 1996). This zinc ion is also believed to be involved in maintenance of the structure. Interestingly, the NS3 proteinase zinc binding site is located in a similar position to that in HRV2-2A^{pro}. As in HRV2-2A^{pro}, the zinc site connects the interbarrel loop (ligands Cys97S^y and Cys99S^y) with the turn between the β strands dII and eII (ligands Cys145S^y and a water

molecule hydrogen bonded to His148). Nevertheless, the binding sites are not identical, as the zinc ion can be removed from the NS3 proteinase simply by dialysis against EDTA (Stempniak *et al.*, 1997). In contrast, removal of the zinc ion from HRV2-2A^{pro} requires prior denaturation of the enzyme with 8 M urea (Sommergruber *et al.*, 1994b; Voss *et al.*, 1995).

Why does HRV2-2A^{pro} possess a zinc binding site? This site in HRV2-2A^{pro} is probably an essential structural element. The reduced size of the N-terminal domain may destabilize the overall fold of the enzyme; the structural zinc site may then compensate by conferring extra stability

Fig. 5. The active site of HRV2-2A^{pro}. Dashed green lines indicate hydrogen bonds; side chains assumed to have a charge are labelled. (A) Superposition of the active sites of the A and B forms. Residues referred to in the text are labelled. (B) Stereo drawing showing the hydrogen-bonding network at the active site of the A form. (C) Hydrogen bonding pattern stabilizing the oxyanion hole of HRV2-2A^{pro}.

on the protein. As HRV2-2A^{pro} is an intracellular enzyme, a zinc binding site might be the only alternative to the introduction of a disulfide bridge. A disulfide bridge in chymotrypsin is found at Cys136–Cys201, which is adjacent to the site that binds zinc in HRV2-2A^{pro} (Tsukada and Blow, 1985).

The HRV2-2A^{pro} active site

The active site of HRV2-2A^{pro} (Figure 5a) consists of a catalytic triad including His18 (the general base), Asp35 and Cys106 (the nucleophile). It is located at the junction of the N- and C-terminal domains (Figure 1A), as is the case for all known chymotrypsin-related serine and cysteine proteinases. Mutagenesis has previously implicated these three amino acids as active site residues in PV1-2A^{pro} (Yu and Lloyd, 1991) and HRV2-2A^{pro} (Sommergruber *et al.*, 1997). In addition to the hydrogen bond to His18N^{δ1}, Asp35 makes hydrogen bonds with Asn16, Tyr86, Thr121 and the main-chain amide nitrogen of His18. Furthermore, Thr121 is involved in a hydrogen bond with Gln91 (Figure 5A and B). All the above-mentioned residues are highly conserved among the enteroviral and rhinoviral 2A cysteine proteinases (Figure 4). Cys106 and His18 at the active site of HRV2-

Table III. Distances (Å) between potential hydrogen donors and acceptors for the active site residues

	A	B
C106S ^γ –H18N ^{ε2}	3.58	3.90
H18N ^{δ1} –D35O ^{δ1}	2.65	2.87
H18N ^{δ1} –D35O ^{δ2}	2.96	3.39
D35O ^{δ2} –N16N ^{δ2}	2.90	2.88
D35O ^{δ2} –T121O ^{γ1}	2.66	2.81
D35O ^{δ1} –Y86O ^η	2.72	2.73
D35O ^{δ1} –H18N	3.02	3.37

2A^{pro} may exist as an imidazolium–thiolate ion pair identical to that found in the papain-like proteinases (E.Bergmann, personal communication).

Figure 5A shows a structural alignment between the active sites of the A and B molecules. Clearly, the conformation of the active site deviates between the two molecules. In particular, a significant variation occurs in the Asp35O^{δ1}–His18N^{δ1} and His18N^{ε2}–Cys106S^γ distances (Table III). Although His18 is well defined in the electron density in both molecules, it probably exhibits conformational variability in solution, as the strand following His18 shows the largest deviation between the two molecules.

A similar flexibility is not observed amongst other chymotrypsin-related serine proteinases, in which the active site is fairly rigid. For example, in α -chymotrypsin, the difference between the two molecules of the asymmetric unit in the distance between the nucleophile (Ser195O^γ) and the general base (His57N^{ε2}) is only 0.1 Å (Tsukada and Blow, 1985; PDB accession code 4CHA). The relatively great flexibility of His18 is probably not crucial for the activity of HRV2-2A^{pro}, given that the active site is a thiolate–imidazolium ion pair. In this case, there are not the same restraints on the relative distance and orientation between the general base and the nucleophile as in the chymotrypsin-like serine proteinases. In these enzymes, such flexibility would most likely be fatal, as a well-established hydrogen bond between donor and acceptor atoms is a prerequisite for proton transfer.

Given the observed flexibility of His18, Asp35 becomes crucial in stabilizing this residue. Asp35 is fixed by an elaborate hydrogen-bonding network (Figure 5A and B). This network involves Asn16, Tyr86 and Thr121 as well as the main-chain amide of His18. Assuming that His18 is an imidazolium ion and Asp35 is a negatively charged carboxylate ion, all hydrogen bonds involving Asp35 have the latter residue as a hydrogen acceptor. In particular, Asn16 and Thr121 are important as hydrogen donors. Mutation of Asn16 to Ala inactivates the enzyme totally (Sommergruber *et al.*, 1997). Thr121 is fully conserved among the enterovirus and rhinoviruses (Figure 4). In addition, replacement of Thr121 in PV1-2A^{pro} with serine had only a marginal effect on activity, indicating the requirement for the hydrogen bond. In contrast, substitution of Thr121 with Asn dramatically affected activity (Yu and Lloyd, 1991). Interestingly, Thr121 is involved in a hydrogen bond with Gln91 acting as a hydrogen donor. Gln91 is also fully conserved among the enterovirus and rhinoviruses (Figure 4). Similarly, Tyr86, which is situated in one of the most flexible regions of HRV2-2A^{pro}, is fully conserved among the enterovirus and rhinoviruses (Figure 4), and also forms a hydrogen bond to Asp35. However, owing to its great flexibility, Tyr86 probably does not stabilize Asp35.

The topology of the active site of HRV2-2A^{pro} resembles that of PV1-3C^{pro} (data not shown). An important difference, however, is that whereas the third member of the catalytic triad is an Asp in HRV2-2A^{pro}, it is a Glu in PV1-3C^{pro} (Mosimann *et al.*, 1997). In HAV-3C^{pro}, the active site consists of a dyad involving the general base (His44) and the cysteine nucleophile. The histidine is stabilized by a hydrogen-bonding network involving a bridging water molecule (Bergmann *et al.*, 1997). Surprisingly, relative to 3C^{pro}, the active site of HRV2-2A^{pro} is more similar to the chymotrypsin-like serine proteinases, all of which have a classical catalytic triad of Ser, His and Asp.

The oxyanion hole

In chymotrypsin-like proteinases, the oxyanion hole serves to stabilize the negative charge developing on the P1-carbonyl oxygen atom during the formation of the tetrahedral intermediate. The oxyanion hole consists of a type II reverse turn preceding the nucleophile and has two main-chain amides pointing towards the P1-carbonyl oxygen atom. The structure of the oxyanion hole in the A molecule of HRV2-2A^{pro} is shown in Figures 5C, 6A

and B. Asp105 forms hydrogen bonds with Glu102 (main-chain N), His63 (main-chain N and N^{δ1}) and Lys62 (main-chain N), with the interaction between Asp105 and His63 being most likely a salt bridge. This hydrogen-bonding network seems to stabilize the oxyanion hole by anchoring the two consecutive type II reverse turns formed by residues 102–108 to a hairpin loop formed by strands aII and bII1 (residues 62–65). The importance of this interaction is shown by experiments in which Asp105 was replaced with either Ser or Asn; both changes inactivated the enzyme (Sommergruber *et al.*, 1997). In addition, Asp105 is also fully conserved among all 2A^{pro} from enterovirus and rhinoviruses, although His63 is not (Figure 4).

The above-mentioned hydrogen-bonding pattern is similar to that in bacterial chymotrypsin-like proteinases such as α -lytic proteinase (α LP; Fujinaga *et al.*, 1985) and SGPB. These chymotrypsin-like serine proteinases have the sequence Gly-Asp-Ser-Gly around the nucleophile, which corresponds to Gly-Asp-Cys-Gly of 2A^{pro}. In all chymotrypsin-like serine proteinases, the Asp forms a hydrogen bond both to the main-chain N at position –3 from the nucleophile in its own reverse turn and to β strands topologically identical to aII and bII1 found in HRV2-2A^{pro} (Figure 5C). In contrast, a similar hydrogen-bond pattern is not found either in PV1-3C^{pro} or in HAV-3C^{pro}. In PV1-3C^{pro}, the residue preceding the nucleophile is Gln, whereas in HAV-3C^{pro} it is Met. In both cases, there are no interactions with strands aII and bII1. Mutation of the active site Cys172 in HAV-3C^{pro} to Ala causes the oxyanion hole to collapse (Allaire *et al.*, 1994). This suggests that the thiolate ion of the proposed imidazolium–thiolate ion pair stabilizes the oxyanion hole by electrostatic interactions, forcing the main-chain carbonyl oxygens from the two residues preceding the nucleophile away from the thiolate ion (E. Bergmann, personal communication). Although we believe that a similar stabilizing imidazolium–thiolate ion pair is present in HRV2-2A^{pro}, it appears that the hydrogen bond network formed by Asp105 confers additional stability on the oxyanion hole. The reason for the differences in hydrogen-bonding patterns between the oxyanion holes of picornaviral 2A^{pro} and 3C^{pro} remains unclear.

Substrate binding

In order to understand the specificity of HRV2-2A^{pro}, an oligopeptide was modelled into the substrate-binding site. As a basis for the modelling, the complex between SGPB and the third domain of the turkey ovomucoid inhibitor (OMTKY3) was used (Read *et al.*, 1983). This complex is a representative of the canonical substrate binding, involving an antiparallel β -strand interaction between P2–P4 from the substrate and the eII strand from the proteinase, which is found in all chymotrypsin-related proteinases (Read and James, 1986; Bode and Huber, 1992). The nomenclature Pn-P1-P1'-Pn' of substrate or inhibitor is that of Schechter and Berger (1967), with the scissile peptide bond lying between P1 and P1'. In addition, the P1-carbonyl oxygen atom points into the oxyanion hole. A strand going from P1' to P5 was modelled into the substrate-binding site in HRV2-2A^{pro} by structurally aligning the oxyanion hole (residues 101–106 in HRV2-2A^{pro} and residues 136–141 in SGPB) and the eII strand (residues 120–124 in HRV2-2A^{pro} and residues 156–159 in

SGPB) between the SGPB-OMKTY3 complex and HRV2-2A^{pro}. Residues 14–19 corresponding to subsites P5–P1' from the OMKTY3 in the SGPB-OMKTY3 complex were then used as a basis for the further modelling of the substrate interactions in HRV2-2A^{pro}. The substrate was manually adjusted to ensure that the P1-carbonyl carbon atom did not form van der Waals overlaps with Cys106 whilst maintaining the antiparallel β -strand interaction. The side chains from the OMKTY3 strand were replaced with the side chains Ile-Ile-Thr-Thr-Ala*Gly (the asterisk marks the cleavage site) corresponding to the sequence around the *cis*-cleavage site between the C terminus of HRV2-VP1 and the N terminus of HRV2-2A^{pro} (Sommergruber *et al.*, 1992). Figure 6A, B and C shows the docking of the substrate to the oxyanion hole and the eII strand. Clearly, severe van der Waals clashes exist between the docked substrate and HRV2-2A^{pro} (Figure 6A), with Tyr85 from HRV2-2A^{pro} forming the most serious overlap with the substrate at P2-Thr. Tyr85 is part of a β -hairpin loop, extending from residue 80 to residue 90, which contains the conserved residues Tyr85, Tyr86 and Pro87. We have named this β -hairpin loop the dityrosine flap.

How can the cleft accommodate a residue at P2 and, furthermore, define the observed preference for Thr? Examination of the differences between the molecules in the asymmetric unit provides a clue. As mentioned, the dityrosine flap is one of the regions showing the largest structural differences between the A and B molecules. The deviation can best be described as a rigid-body rotational movement of the loop around two hinges formed by the peptide bonds from residues 74–75 and 94–95 in HRV2-2A^{pro} (Figure 1B). Residues 82–85 from this β -hairpin loop form one side of the substrate-binding cleft. Owing to the difference in conformation of the dityrosine flap, the width of the substrate-binding cleft differs by 1–1.5 Å between the two molecules in the asymmetric unit. The cleft is narrower in the A molecule, in which severe van der Waals clashes are found not only between the modelled oligopeptide and Tyr85, but also between the modelled side-chain atoms of the oligopeptide and the main-chain atoms of all residues from 83 to 85. In the B molecule, however, the only van der Waals clashes are those between Tyr85 from HRV2-2A^{pro} and P2-Thr. This structural difference may be representative of the dynamic behaviour of the dityrosine flap in solution. In addition, the fairly long minimal substrate spanning P7 to P2' (Sommergruber *et al.*, 1992) may be required to provide sufficient binding energy to overcome the partial occlusion of the binding cleft by this portion of the dityrosine flap. As the B molecule of HRV2-2A^{pro} contains the most open substrate-binding cleft with the least van der Waals clashes, only this model will be discussed further.

A prerequisite for binding of the Thr residue at P2 is the rotation of Tyr85 around its C α –C β bond, away from its original position (Figure 6B). Interestingly, in this model, P2-Thr is able to form a hydrogen bond with Ser83

from the dityrosine flap. Ser83 is not a conserved residue; however, all residues in the equivalent position of other 2A proteinases have potential as hydrogen bond acceptor or donor atoms (Figure 4). In the free enzyme, the side chain of Tyr85 is stacked upon His18. The ability of P2-Thr to form a hydrogen bond with Ser83 from HRV2-2A^{pro} might compensate for turning Tyr85 away from its stacked position. In addition, the C γ^2 methyl group from the P2-Thr residue might provide further favourable binding energy in terms of additional buried accessible surface area. These features could explain why peptide substrates with a Thr in the P2 position are accepted 5-fold more efficiently than those with Ser at P2 and why no cleavage is observed for those peptides bearing a Val at P2 (Sommergruber *et al.*, 1992).

The model of substrate binding also provides explanations for the absolute requirement of a Gly in the P1' position and for the strong preference for either a Leu or Ile in the P4 position. Leu19 from HRV2-2A^{pro} blocks any access to a possible S1' site (Figure 6B). In contrast, the S4 site is a fairly narrow hydrophobic pocket that favours the reasonably flexible hydrophobic side chains in residues such as Leu or Ile (Figure 6C).

The S1 site

The wild-type amino acid at P1 in the HRV2 viral polyprotein cleavage site is alanine. However, oligopeptide substrates having the residues Thr, Leu, Phe and Tyr at P1 have comparable cleavage efficiencies, as measured by k_{cat}/K_m values (Sommergruber *et al.*, 1992). In addition, an oligopeptide with Met at P1 was cleaved 5-fold better than the wild-type peptide. Furthermore, substrates containing the positively charged residues Arg and Lys were also cleaved, albeit less efficiently, whereas substrates containing Asp and Glu at P1 were not cleaved (Sommergruber *et al.*, 1992).

How does the model of substrate binding to HRV2-2A^{pro} account for these findings? The surface of the S1 site is shown in Figure 6C. The S1 pocket, formed by residues from the eII strand (residues 122–123) and the oxyanion hole (residues 101–106), is fairly narrow and flat and filled out mainly by the side chain of Cys101. This residue is not conserved among the entero- and rhinoviruses, but is replaced with either a Ser or an Ala (Figure 4). Based on this flat and narrow character, the long and flexible methionine side chain could easily be accommodated in the S1 site. Side chains such as Phe and Tyr do not have the same conformational freedom. It is very likely that the favourable binding energy gained by having a large buried surface area is outweighed by having to force the side chain into an unfavourable conformation. This provides a reasonable explanation of why peptides with these side chains have similar cleavage efficiencies to those with Ala. At the bottom of the pocket, a negatively charged region is provided by the carboxylate side chain of Glu102. This negative charge probably prevents P1 Glu and Asp side chains from entering the

Fig. 6. A model of substrate binding in HRV2-2A^{pro}. The main chain of the modelled substrate is shown in purple. (A) Model of the substrate binding in the substrate-binding cleft of the A form (red, main chain) and the B form (blue, main chain). Side chains referred to in the text are labelled. (B) Stereo drawing showing the oligopeptide substrate model bound to the HRV2-2A^{pro} substrate-binding cleft (white, main chain) in the B molecule. Tyr85 has been rotated in order to form an S2 site [compare (A)]. (C) Stereo drawing of the surface of the substrate-binding cleft in HRV2-2A^{pro}. Part of the dityrosine flap (residues 82–89) has been removed. This panel was generated by VMD (Humphrey *et al.*, 1996).

S1 pocket due to unfavourable electrostatic interactions, whilst supporting substrates with P1 Lys or Arg, such as that found in the eIF4G cleavage site utilized by the virus in the shut-off of host-cell protein synthesis (Lamphear *et al.*, 1993).

HRV2-2A^{pro} self-processing: a *cis*-event?

It has been proposed that HRV2-2A^{pro} performs the first cleavage of the polyprotein as an intramolecular (*cis*) cleavage between its own N terminus and the C terminus of the capsid protein, VP1, while still on the ribosome (Toyoda *et al.*, 1986). However, recent evidence from IRES screening experiments indicates that the proteolytic processing of the polyprotein by 2A^{pro} might not occur co-translationally (Paul *et al.*, 1998). Nevertheless, the *cis* cleavage constitutes a special case because the P1' residue is at the same time the N-terminal residue of 2A^{pro} so that the substrate is covalently attached to the proteinase. In all characterized viral proteinase *cis*-cleavage sites, the P1' residue is a Gly; this appears to be a general requirement for both *cis* and *trans* cleavage. However, the specificity requirements for the other sites (P4 and P2) may not be as stringent for the *cis* cleavage. In a study by Hellen *et al.* (1992), it was found that the specificity requirement was less stringent for the *cis* cleavage than for the *trans* cleavage for PV1-2A^{pro}. As the substrate is covalently attached to 2A^{pro} before the *cis* cleavage, this huge favourable entropic term for substrate binding might explain the less stringent specificity.

In the case of HRV2-2A^{pro}, the oligopeptide derived from the *cis*-cleavage site between HRV2-VP1 and HRV2-2A^{pro} (Thr-Arg-Pro-Ile-Ile-Thr-Thr-Ala*Gly-Pro-Ser-Asp-Met) is also a good *trans* substrate of HRV2-2A^{pro} (Sommergruber *et al.*, 1992). In the HRV2-2A^{pro} structure, the N terminus forms a type I reverse turn. The well-defined electron density and low *B*-factors in both molecules in the asymmetric unit suggest that this loop is a very stable part of the HRV2-2A^{pro} structure. A strong hydrogen bond is observed between the protonated α -amino group of Gly1 and the carboxylate side chain of Asp4 (Figure 7A). However, this is not a conserved motif among 2A proteinases from either rhino- or enteroviruses (Figure 4). It is, therefore, difficult to predict whether the loop is a conserved structural feature of all the 2A cysteine proteinases and whether it serves as a general protection against product inhibition. The 3C proteinases of both entero- and rhinoviruses also have *cis*-cleavage activity. In the known 3C^{pro} structures of PV1, HRV14 and HAV, the N-terminal residues are helical, a structure that has been proposed to protect the enzyme from product inhibition (Matthews *et al.*, 1994; Bergmann *et al.*, 1997; Mosimann *et al.*, 1997). No such helix is present in HRV2-2A^{pro}.

The structure of HRV2-2A^{pro} provides a picture of the enzyme after *cis* cleavage has taken place. Immediately prior to the reaction, however, the N terminus of 2A^{pro} and the C terminus of VP1 must have been covalently linked together at the active site. Thus, Gly1 must be bound at the S1' site during the *cis* cleavage. Inspection of the structure of HRV2-2A^{pro} shows that rearrangement of HRV2-2A beyond Met5 would probably affect the stability of the N-terminal β sheet and thus also of the active site. It is, however, indeed possible to bring the N terminus in proximity to the S1' site by adjusting only

Fig. 7. The N terminus of HRV2-2A^{pro}. (A) The strong hydrogen bond at the type I reverse turn at the N terminus. (B) The position of the N terminus in the free enzyme is shown thick and blue. The assumed position and conformation of the N terminus prior to *cis* cleavage, in which Gly1 is situated at the P1' site, is thick and red. Met5 is green.

the main chain and torsion angles of the first five residues (Figure 7B). The 2A cysteine proteinase must therefore be otherwise fully folded prior to *cis* cleavage.

Interestingly, the *cis* cleavage between VP1 and 2A^{pro} resembles the processing of the proenzyme of α LP (Sauter *et al.*, 1998). α LP is a chymotrypsin-like bacterial serine proteinase consisting of 198 residues. The proenzyme of α LP contains an N-terminal extension (166 residues) called the prosegment that acts as a foldase catalysing the folding of an intermediate state of α LP to the active, mature enzyme. Subsequently, the proenzyme is processed by a *cis* cleavage between the C terminus of the prosegment and the N terminus of α LP. Recently, the crystal structure

of a stable complex between the free prosegment and mature α LP was solved (Sauter *et al.*, 1998), revealing that the C terminus of the prosegment was bound in a product-inhibited fashion to the substrate-binding cleft (subsites S1–S4). The N terminus of α LP in the α LP–prosegment complex occupies the same position observed in the native enzyme (Fujinaga *et al.*, 1985), as part of the end of a β strand distant from the S1' site. Thus, both the prosegment and α LP must become fully folded prior to the *cis*-cleavage event; furthermore, only the N terminus of the α LP must rearrange. It is very likely that a similar process occurs for the *cis* cleavage between VP1 and 2A^{pro} in which the two structures must be fully folded in their native states except for the C terminus of VP1, which occupies subsites from S7 to S1 and the N terminus of 2A^{pro}.

HRV2-2A^{pro} cleavage of eIF4G

The actual mechanism of cleavage of eIF4G by 2A^{pro} leading to the shut-down of the host cell's translational machinery continues to be the subject of much debate (Liebig *et al.*, 1993; Bovee *et al.*, 1998 and references therein). Two mechanisms have been proposed: a direct cleavage of eIF4G by 2A^{pro} or 2A^{pro} activation of a latent cellular proteinase that is then responsible for the cleavage of eIF4G. Although the weight of evidence supports the direct cleavage of eIF4G by 2A^{pro}, the poor cleavage of pure eIF4G *in vitro* has raised doubts about the direct cleavage. The structure of HRV2-2A^{pro} supports a direct cleavage event. The amino acids recognized by HRV2-2A^{pro} at its cleavage site on eIF4G (Leu-Ser-Thr-Arg*Gly-Pro; Lamphear *et al.*, 1993; Sommergruber *et al.*, 1994a) can be accommodated at the active site. In addition, studies of the cleavage mechanism of eIF4G by HRV2-2A^{pro} and PV1-2A^{pro} have shown that the cleavage of eIF4G is improved significantly when it is part of a complex with eIF4E, as in the eIF4F complex (Haghighat *et al.*, 1996; Ventoso *et al.*, 1998). Given the requirements of a long peptide substrate and a fairly long and narrow substrate-binding groove of HRV2-2A (Figure 6C), it is plausible that HRV2-2A^{pro} has evolved to recognize a conformation of eIF4G that is formed when eIF4E is bound, thus explaining the higher cleavage efficiency of eIF4G in this state. The structure of HRV2-2A^{pro} would thus support a direct cleavage mechanism of eIF4G as part of eIF4F.

Why is a second proteinase necessary for entero- and rhinoviruses?

The similarity of the HRV2-2A^{pro} and the picornaviral 3C^{pro} folds implies that the two cysteine proteinases have arisen by gene duplication. Why have the entero- and rhinoviruses evolved two cysteine proteinases having similar catalytic mechanisms based on cysteine nucleophiles but with different specificity? Obviously, the viruses benefit from the presence of this additional proteinase. Although 2A^{pro} has been suggested to take part in other processes such as RNA replication, we believe that one of the primary benefits of 2A^{pro} is in promoting its own IRES by causing the host-cell translational shut-off. As both entero- and rhinoviruses have a weak-binding IRES compared with the other picornaviruses, the processing of eIF4G by 2A^{pro} is most likely vital in order for the IRES

to compete with the host cell's cap-binding complex. In the absence of a canonical cleavage site for a 3C^{pro} in the region of eIF4G between the eIF4E and eIF3 binding sites, it seems likely that further evolution of 3C^{pro} to perform the eIF4G cleavage would not be possible without affecting polyprotein processing and RNA binding. eIF4G is also cleaved during replication of foot-and-mouth disease virus; however, 3C^{pro} is once again not responsible for this reaction. Instead, a papain-like cysteine proteinase, designated the leader proteinase, performs this cleavage (Devaney *et al.*, 1988) at a site six amino acids upstream from that of 2A^{pro} cleavage (Kirchweiger *et al.*, 1994; Guarné *et al.*, 1998), also supporting the idea that 3C^{pro} cannot evolve to carry out the reaction required to induce the host-cell shut-off. Interestingly, recent studies indicate that 2A^{pro} is also involved in the cleavage of the poly(A)-binding protein (PABP) (Joachims *et al.*, 1999; Kerekatte *et al.*, 1999). PABP is believed to be important for initiation of the host cell's translation and has been shown to interact with eIF4G (Sachs *et al.*, 1997). This further indicates that the proteolytic *trans* cleavage activity of 2A^{pro} is vital for entero- and rhinoviruses in promoting their weak-binding IRES. Indeed, the commitment of precious coding space to an additional proteinase for the cleavage of eIF4G and PABP emphasizes the importance of this reaction for viral replication. The structure of HRV2-2A^{pro} presented here represents the first views of a cysteine proteinase with a chymotrypsin-like fold that is capable of inducing picornaviral host-cell shut-off.

Materials and methods

Crystallization

Expression and purification of HRV2-2A^{pro} were as reported previously (Liebig *et al.*, 1993). HRV2-2A^{pro} was dissolved in 10 mM β -mercaptoethanol at a concentration of 20 mg/ml. Crystals were grown at 4°C by the hanging drop vapour diffusion method by mixing 2 μ l of the protein solution and 2 μ l of crystallization buffer, and equilibrating the drop against the crystallization buffer (1 ml). The crystallization buffer consisted of 7.5% PEG8K, 0.3 M KCNS, tertiary amyl alcohol 5% v/v, glycerol 10% v/v, 10 mM β -mercaptoethanol and 50 mM sodium potassium phosphate pH 8.0. Crystals grew within 1–2 weeks and had typical dimensions of $0.5 \times 0.08 \times 0.08$ mm³.

Data collection

Two datasets, one for MAD phasing and a second high-resolution dataset to be used for refinement, were collected at the National Synchrotron Light Source, Brookhaven National Laboratory. All data were collected on flash-frozen crystals at 100 K by transferring the crystal rapidly into a cryoprotectant containing mother liquor made up to 30% v/v glycerol. The MAD data were collected on beamline X12C, equipped with a Brandeis 10×10 cm² CCD detector. The MAD experiment was conducted as a classical three-wavelength experiment over the absorption edge of zinc. Two of the wavelengths, L1 and L2, were chosen at the inflection point and the peak of the absorption edge of zinc (1.2835 and 1.2827 Å, respectively), whereas the last (L3) was a remote wavelength at 1.2450 Å. The data at each wavelength were collected in a single sweep through reciprocal space. The high-resolution data were collected on beamline X25 using a 20×20 cm² Brandeis CCD detector. The wavelength was set at 1.1 Å. All images collected were integrated and scaled using the programs DENZO and SCALEPACK from the HKL package (Otwinowski and Minor, 1997). All data were scaled using the Wilson plot program Truncate in the CCP4 program suite (Collaborative Computational Project Number 4, 1994).

Phasing and refinement

The MAD data collected from X12C were treated as three derivatives having anomalous scattering contributions. L1 was chosen as the reference data contributing with an anomalous signal only, whereas L2

and L3 both contributed with a dispersive and anomalous signal. All three datasets were put on a common scale using the CCP4 program SCALEIT. The positions of the zinc atoms were found by inspecting the anomalous Patterson map computed from dataset L2. SHELXS-90 (Sheldrick, 1991) was used in interpreting the Patterson map. The two Zn positions and their dispersive and anomalous occupancies were refined using the program MLPHARE (Otwinowski, 1991). The electron density maps from the phase determination showed clear boundaries between protein and solvent channels. Further density modification was carried out in two steps using the CCP4 program DM (Cowtan and Main, 1996). Initially, only solvent flattening and histogram matching were applied to 2.7 Å. The resulting electron density map clearly revealed the positions of the two molecules in the asymmetric unit and the NCS operator was determined by manually aligning the bones skeleton of one of the molecules in the asymmetric unit to superimpose on top of the other. In the second step, solvent flattening, histogram matching and 2-fold averaging were performed; the phases were extended to 2.2 Å. The resulting electron density was of excellent quality and the entire peptide chain in both molecules, except for the last three residues at the C terminus, could be traced unambiguously.

The initial model was built using TURBO-FRODO based on the 2.2 Å resolution electron density map. The model was subsequently refined against the 1.95 Å resolution data using simulated annealing and individually restrained *B*-factor refinement with the program CNS (Brunger *et al.*, 1998) including all measured reflections. Both overall anisotropic *B*-factor corrections and bulk solvent corrections were applied. In the early stages of the refinement, NCS restraints were applied, but were relaxed in the later stages without any dramatic change in either the working *R*-factor or in the free *R*-factor. In the last steps of the refinement, water molecules were included only if there was spherical density in both the $F_o - F_c$ and $2F_o - F_c$ electron density maps, and the peaks were within hydrogen-bonding distance of a hydrogen bond donor or acceptor atoms in the protein molecules. Finally, all water molecules having a *B*-factor of $>70 \text{ Å}^2$ after *B*-factor refinement were removed from the coordinate list. The atomic coordinates have been deposited in the Protein Data Bank, accession code 2HRV.

Acknowledgements

We wish to thank Robert M. Sweet at NSLS, Brookhaven National Laboratory, for his assistance and advice during the data collection, Marie Fraser for her assistance during the data collection at BNL, and Dragana Jugovic and Martin Langer for technical assistance. This work was supported by a group grant from the Canadian MRC to the MRC group in Protein Structure and Function at the University of Alberta (to M.N.G.J.), by grants P12193-MOB and F508-MED from the Austrian Science Foundation (to E.K.), and by the Josefine Hirtl Stiftung for medical research (to T.S.). J.F.W.P. was supported by a Killam postdoctoral fellowship.

References

Alberts, I.L., Nadassy, K. and Wodak, S.J. (1998) Analysis of zinc binding sites in protein crystal structures. *Protein Sci.*, **17**, 1700–1716.

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

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. *Nature Med.*, **3**, 320–326.

Bazan, J.F. and Fletterick, R.J. (1988) Viral cysteine proteinases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. *Proc. Natl Acad. Sci. USA*, **85**, 7872–7876.

Belsham, G.J. and Sonenberg, N. (1996) RNA–protein interactions in regulation of picornavirus RNA translation. *Microbiol. Rev.*, **60**, 499–511.

Bergmann, E.M., Mosimann, S.C., Chernaia, M.M., Malcolm, B.A. and James, M.N.G. (1997) The refined crystal structure of the 3C gene product from hepatitis A virus: Specific proteinase activity and RNA recognition. *J. Virol.*, **71**, 2436–2448.

Bode, W. and Huber, R. (1992) Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.*, **201**, 433–451.

Borman, A.M., Kirchweiger, R., Ziegler, E., Rhoads, R.E., Skern, T. and Kean, K. (1997) Role of eIF-4γ proteolysis on the internal initiation of protein synthesis. *RNA*, **3**, 186–196.

Bovee, M.L., Marissen, W.E., Zamora, M. and Lloyd, R.E. (1998) The predominant eIF4G-specific cleavage activity in poliovirus-infected HeLa cells is distinct from 2A protease. *Virology*, **245**, 229–240.

Brunger, A.T. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.

Choi, H., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M.G. and Wengler, G. (1991) Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature*, **354**, 37–43.

Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.

Cowtan, K.D. and Main, P. (1996) Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallogr. D*, **52**, 43–48.

Devaney, M.A., Vakharia, V.N., Lloyd, R.E., Ehrenfeld, E. and Grubman, M.J. (1988) Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. *J. Virol.*, **62**, 4407–4409.

Dougherty, W.G. and Semler, B.L. (1993) Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiol. Rev.*, **57**, 781–822.

Fujinaga, M., Delbaere, L.T.J., Brayer, G.D. and James, M.N.G. (1985) Refined structure of α-lytic protease at 1.7 Å resolution. Analysis of hydrogen bonding and solvent structure. *J. Mol. Biol.*, **183**, 479–502.

Gorbalenya, A.E., Donchenko, A.P., Blinov, V.M. and Koonin, E.V. (1989) Cysteine proteinases of positive strand RNA viruses and chymotrypsin-like serine proteinases. *FEBS Lett.*, **243**, 103–114.

Gradi, A., Svitkin, Y.V., Imataka, H. and Sonenberg, N. (1998) Proteolysis of human eIF4GII, but not eIF4GI, coincided with the shutoff of host protein synthesis after poliovirus infection. *Proc. Natl Sci. Acad. USA*, **95**, 11089–11094.

Guarné, A., Tormo, J., Kirchweiger, R., Pfistermueller, D., Fita, I. and Skern, T. (1998) Structure of the foot-and-mouth disease virus leader protease: a papain-like fold adapted for self-processing and eIF4G recognition. *EMBO J.*, **17**, 7469–7479.

Haghighat, A., Svitkin, Y., Novoa, I., Kuechler, E., Skern, T. and Sonenberg, N. (1996) The eIF4G–eIF4E complex is the target for direct cleavage by the rhinovirus 2A proteinase. *J. Virol.*, **70**, 8444–8450.

Hambidge, S.J. and Sarnow, P. (1992) Translational enhancement of the poliovirus 5′ noncoding region mediated by the virus-encoded polypeptide 2A. *Proc. Natl Acad. Sci. USA*, **89**, 10272–10276.

Hellen, C.U.T., Lee, C.-K. and Wimmer, E. (1992) Determinants of substrate recognition by poliovirus 2A proteinase. *J. Virol.*, **66**, 3330–3338.

Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD–visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.

Hutchinson, E.G. and Thornton, J.M. (1996) PROMOTIF—A program to identify and analyze structural motifs in proteins. *Protein Sci.*, **5**, 212–220.

Jackson, R.J., Howell, M.T. and Kaminski, A. (1990) The novel mechanism of initiation of picornavirus RNA translation. *Trends Biochem. Sci.*, **15**, 477–483.

Jang, S.K., Pestova, T.V., Hellen, C.U.T., Witherell, G.W. and Wimmer, E. (1990) Cap-independent translation of picornavirus RNAs: structure and function of the internal ribosomal entry site. *Enzyme*, **44**, 292–309.

Joachims, M., Breugel, P.C.V. 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.

Kerekatte, V., Keiper, B.D., Badorff, C., Cai, A., Knowlton, K.U. and Rhoads, R.E. (1999) Cleavage of poly(A)-binding protein by Coxsackievirus 2A protease *in vitro* and *in vivo*: Another mechanism for host protein synthesis shutoff? *J. Virol.*, **73**, 709–717.

Kim, J.L. *et al.* (1996) Crystal structure of the hepatitis C virus NS3 proteinase domain complexed with a synthetic NS4A cofactor peptide. *Cell*, **87**, 343–355.

Kirchweiger, R. *et al.* (1994) Foot-and-mouth disease virus leader protease: purification of the Lb form and determination of its cleavage site on eIF4γ. *J. Virol.*, **68**, 5677–5684.

Kraulis, P.J. (1991) MOLSCRIPT—a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Lamphear, B.J., Yan, R., Yang, D., Waters, D., Liebig, H.-D., Klump, H., Kuechler, E., Skern, T. and Rhoads, R.E. (1993) Mapping the cleavage site in protein synthesis initiation factor eIF-4G of the 2A proteases from human coxsackievirus and rhinovirus. *J. Biol. Chem.*, **268**, 19200–19203.

- Liebig,H.-D. *et al.* (1993) Purification of two picornaviral 2A proteinases: Interaction with eIF-4G and influence on *in vitro* translation. *Biochemistry*, **32**, 7581–7588.
- Lloyd,R.E., Grubman,M. and Ehrenfeld E. (1988) Relationship of p220 cleavage during picornavirus infection to 2A sequences. *J. Virol.*, **62**, 4216–4223.
- Love,A.L., Parge,H.E., Wickersham,J.A., Hostomsky,Z., Habuka,N., Moomaw,E.W., Adachi,T. and Hostomska,Z. (1996) The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell*, **87**, 331–342.
- Lu,H.H., Li,X., Cuconati,A. and Wimmer,E. (1995) Analysis of picornavirus 2A^{pro} proteins: Separation of proteinase from translation and replication functions. *J. Virol.*, **69**, 7445–7452.
- Matthews,B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.*, **33**, 491–497.
- Matthews,D.A. *et al.* (1994) Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA binding site and means for cleaving precursor polypeptide. *Cell*, **77**, 761–771.
- Merrit,E.A. and Bacon,D.J. (1997) Raster3D: photorealistic molecular graphics. *Methods Enzymol.*, **277**, 505–524.
- Molla,A., Paul,A.V., Schmid,M., Jang,S.K. and Wimmer,E. (1993) Studies on dicistronic polioviruses implicate viral proteinases 2A^{pro} in RNA replication. *Virology*, **196**, 739–747.
- Mosimann,S.C., Cherney,M.M., Sia,S., Plotch,S. and James,M.N.G. (1997) Refined X-ray crystallographic structure of the poliovirus 3C gene product. *J. Mol. Biol.*, **273**, 1032–1047.
- Otwinowski,Z. (1991) Maximum likelihood refinement of heavy atom parameters. In Wolf,W., Evans,P.R. and Leslie,A.G.W. (eds), *Proceedings of the CCP4 Study Weekend*. pp. 80–86.
- Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **277**, 307–325.
- Paul,A.V., Mugavero,J., Molla,A. and Wimmer,E. (1998) Internal ribosomal entry site scanning of poliovirus polypeptide: implications for proteolytic processing. *Virology*, **250**, 241–253.
- Read,R.J. and James,M.N.J. (1986) Introduction to the protein inhibitors: X-ray crystallography. In Barrett,A.J. and Salvesen,G. (eds), *Proteinase Inhibitors*. Elsevier, Amsterdam, The Netherlands, pp. 301–336.
- Read,R.J., Fujinaga,M., Sielecki,A.R. and James,M.N.G. (1983) Structure of the complex of *Streptomyces griseus* protease B and the third domain of the turkey ovomucoid inhibitor at 1.8-Å resolution. *Biochemistry*, **22**, 4420–4433.
- Ryan,M. and Flint,M. (1997) Virus-encoded proteinases of the picornavirus super-group. *J. Gen. Virol.*, **78**, 699–723.
- Sachs,A.B., Sarnow,P. and Hentze,M.W. (1997) Starting at the beginning, middle and end: translation initiation in eukaryotes. *Cell*, **89**, 831–838.
- Sauter,N.K., Mau,T., Rader,S.D. and Agard,D.A. (1998) Structure of α -lytic protease complexed with its pro region. *Nature Struct. Biol.*, **5**, 945–950.
- Schechter,I. and Berger,A. (1967) On the size of the active site in proteases I. Papain. *Biochem. Biophys. Res. Commun.*, **27**, 157–162.
- Sheldrick,G.M. (1991) Heavy atom location using SHELXS-90. In Wolf,W., Evans,P.R. and Leslie,A.G.W. (eds), *Proceedings of the CCP4 Study Weekend*. pp. 23–38.
- Sommergruber,W. *et al.* (1992) Cleavage specificity on synthetic peptide substrates of human rhinovirus 2 proteinase 2A. *J. Biol. Chem.*, **31**, 22639–22644.
- Sommergruber,W. *et al.* (1994a) 2A proteinases of coxsackie- and rhinovirus cleave peptides derived from eIF-4G via a common recognition motif. *Virology*, **198**, 741–745.
- Sommergruber,W., Casari,C., Fessl,F., Seipelt,J. and Skern,T. (1994b) The 2A proteinase of human rhinovirus is a zinc containing enzyme. *Virology*, **204**, 815–818.
- Sommergruber,W., Seipelt,J., Fessl,F., Skern,T., Liebig,H.-D. and Casari,G. (1997) Mutational analyses support a model for the HRV2-2A proteinase. *Virology*, **234**, 203–214.
- Stempniak,M., Hostomska,Z., Nodes,B.R. and Hostomsky,Z. (1997) The NS3 proteinase domain of hepatitis C virus is a zinc-containing enzyme. *J. Virol.*, **71**, 2881–2886.
- Toyoda,H., Nicklin,M.J.H., 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 polypeptide. *Cell*, **45**, 761–770.
- Tsukada,H. and Blow,D.M. (1985) Structure of α -chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol.*, **184**, 703–711.
- Ventoso,I., McMillan,S.E., Hershey,J.W.B. and Carrasco,L. (1998) Poliovirus 2A proteinase cleaves directly the eIF-4G subunit of eIF-4F complex. *FEBS Lett.*, **435**, 79–83.
- Voss,T., Meyer,R. and Sommergruber,W. (1995) Spectroscopic characterisation of rhinoviral 2A: Zn is essential for the structural integrity. *Protein Sci.*, **4**, 2526–2531.
- Yu,S.F. and Lloyd,R.E. (1991) Identification of essential amino acid residues in the functional activity of poliovirus 2A protease. *Virology*, **182**, 615–625.
- Yu,S.F., Benton,P., Bovee,M., Sessions,J. and Lloyd,R.E. (1995) Defective RNA replication by poliovirus mutants deficient in 2A protease cleavage activity. *J. Virol.*, **69**, 247–252.
- Ziegler,E., Borman,A.M., Deliat,F.G., Liebig,H.-D., Jugovic,D., Kean,K.M., Skern,T. and Kuechler,E. (1995) Picornavirus 2A proteinase-mediated stimulation of internal initiation of translation is dependent of enzymatic activity and the cleavage products of cellular proteins. *Virology*, **213**, 549–557.

Received July 2, 1999; revised August 5, 1999;
accepted August 24, 1999