Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Molecules in focus

Foot-and-mouth disease virus 3C protease: Recent structural and functional insights into an antiviral target

Stephen Curry a,*, Núria Roqué-Rosell b, Patricia A. Zunszain a, Robin J. Leatherbarrow b

a Biophysics Section, Division of Cell and Molecular Biology, Blackett Laboratory, Imperial College, Exhibition Road, London SW7 2AZ, UK
b Biological and Biophysical Chemistry Section, Department of Chemistry, Imperial College, Exhibition Road, London SW7 2AZ, UK

Received 6 July 2006; received in revised form 25 July 2006; accepted 26 July 2006
Available online 14 August 2006

Abstract

The 3C protease from foot-and-mouth disease virus (FMDV 3Cpro) is critical for viral pathogenesis, having vital roles in both the processing of the polyprotein precursor and RNA replication. Although recent structural and functional studies have revealed new insights into the mechanism and function of the enzyme, key questions remain that must be addressed before the potential of FMDV 3Cpro as an antiviral drug target can be realised.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Foot-and-mouth disease virus; 3C protease; Protease mechanism; Peptide cleavage specificity; Anti-viral drugs

1. Introduction

FMDV is the causative agent of an extraordinarily contagious disease of cloved-hooved animals such as cattle, sheep, pigs and goats. Although not usually fatal to the infected animal, the speed of disease transmission – largely through the inhalation of virus aerosols – means that drastic measures must be deployed to control outbreaks (Grubman & Baxt, 2004). Economically devastating epidemics continue to occur across the globe, in part because of ongoing technical and political challenges associated with the use of FMDV vaccines. Alternative control mechanisms, based on an understanding of the molecular aspects of viral replication and pathogenesis, are therefore being eagerly sought.

FMDV is a member of the aphthovirus genus of the picornavirus family, an important group of mammalian single-strand, positive-sense RNA viruses that includes poliovirus (PV), human rhinovirus (HRV) and hepatitis A virus (HAV) (Mason, Grubman, & Baxt, 2003). Picornaviruses share a common replication strategy requiring the translation of a polyprotein precursor that is cleaved by virally encoded proteases into the proteins that make up the viral capsid and replication machinery (Leong, Cornell, & Semler, 2002; Mason et al., 2003). The conserved 3C protease, the only picornaviral protease common to all genera, is a key player in this process and is a potential anti-viral drug target. Recent investigations have unearthed new details about the proteolytic
and other roles of FMDV 3C\textsuperscript{pro} and will be the focus of this review.

2. Expression of FMDV 3C\textsuperscript{pro} during infection

The FMDV infection process begins once the virus has delivered the viral RNA to the cytoplasm of the host cell. The RNA serves immediately as an mRNA to direct synthesis of a polyprotein of over 2300 amino acids which is ultimately cleaved into fourteen separate proteins; ten of the thirteen cleavages are performed by FMDV 3C\textsuperscript{pro} which is initially contained within the viral polyprotein (Fig. 1) (Mason et al., 2003).

3. Structure and proteolytic mechanism

The crystal structure of FMDV 3C\textsuperscript{pro} confirmed that, in common with other picornaviral 3C proteases, it adopts a fold similar to that of the archetypal serine protease chymotrypsin (Birtley et al., 2005) and belongs to an unusual family of chymotrypsin-like cysteine proteases (Barrett & Rawlings, 2001). The protein comprises two six-strand β-barrels connected by a short linker that pack together with the barrel axes at approximately 90° to one another; between these barrels on one face of the protein lies the peptide binding cleft that also accommodates the active site of the enzyme (Fig. 2a). The initial crystal structure suggested that a two-strand β-ribbon structure, which is observed to fold over the peptide binding cleft in other picornavirus 3C proteases (Bergmann et al., 1995; Matthews et al., 1994; Mosimann, Cherney, Sia, Plotch, & James, 1997), was disordered in FMDV 3C\textsuperscript{pro}. However, a more recent structure determination, using a mutant designed to generate a different crystal form, reveals that this β-ribbon is indeed conserved in FMDV 3C\textsuperscript{pro} (Sweeney et al., 2006). Although the β-ribbon (residues 138–150) clearly possesses a degree of flexibility, it makes important contributions to substrate specificity via direct contacts with peptide sequences bound in the active site cleft (Matthews et al., 1999). Indeed, this loop probably serves to position the substrate appropriately for proteolysis since mutagenesis of Cys\textsubscript{42} at the apical tip of the β-ribbon was shown to have a significant impact on catalytic activity. The hydrophobic nature of this amino acid appears to be of particular importance: the C142S substitution reduced activity by two orders of magnitude, whereas the modification C142L resulted in essentially wild-type activity (Sweeney et al., 2006).

3.1. Catalytic triad and mechanism

The catalytic mechanism of picornaviral 3C proteases is an enduring puzzle (Skern et al., 2002). Although sequence alignments indicated that 3C proteases had a Cys-His-Asp/Glu catalytic triad akin to the Ser-His-Asp triad found in serine proteases, early crystallographic work highlighted intriguing differences between HRV and HAV 3C\textsuperscript{pro} on the one hand and chymotrypsin-like serine proteases on the other. In particular, the structures showed that the third member of the triad (Glu in HRV; Asp in HAV) did not interact appropriately with the catalytic His residue and therefore suggested that it had a lesser or non-existent role in catalysis than is the case for serine proteases (Allaire, Cherney, Malcolm, & James, 1994; Bergmann et al., 1997; Matthews et al., 1994). However, the latest structural work on FMDV and HAV 3C\textsuperscript{pro} (Birtley et al., 2005; Sweeney et al., 2006; Yin et al., 2005) and on related 3C-like viral proteases (Phan et al., 2002; Zeitler, Estes, & Prasad, 2006) helps to establish that this class of enzymes all possess a Cys-His-Asp/Glu triad in the active site that is very similar in conformation to the Ser-His-Asp triad found in chymotrypsin-like serine proteases (Hedstrom, 2002) (Fig. 2b). These structural observations argue strongly in favour of a significant role for the third member of the triad in picornaviral 3C proteases and are consistent both with its strict conservation as Asp or Glu in 3C\textsuperscript{pro} sequences and the observation that substitution of this residue is invariably severely detrimental to catalytic activity (Grubman, Zellner, Bablanian, Mason, & Piccone, 1995; Kean, Teterina, Marc, & Girard, 1991).

However, despite structural similarities and an evident evolutionary relationship with chymotrypsin-like serine proteases (Barrett & Rawlings, 2001; Brenner, 1988), the precise mechanistic details of 3C\textsuperscript{pro} have yet to be elucidated. Biochemical studies have found that PV 3C\textsuperscript{pro} does not possess a Cys-His thiolate-imidazolium ion pair in the active site and have therefore ruled out a catalytic mechanism similar to papain-like cysteine pro-
teases (Sarkany, Szeltner, & Polgar, 2001). Nevertheless, definitive evidence for a general base catalytic mechanism, as found for chymotrypsin-like serine proteases, has yet to be obtained (Sarkany & Polgar, 2003). It is also still unclear why some picornavirus 3C proteases have Glu as the third member of the triad when this residue is almost invariably an Asp in serine proteases (Sarkany et al., 2001). An intriguing outlying observation is that the 3C-like main protease from coronavirus only contains a Cys-His pair in the active site; the position normally occupied by Asp is taken by a hydrophobic residue (Cys or Val) (Anand et al., 2002). Clearly further work will be required to make a definitive determination of the proteolytic mechanism of FMDV 3C<sub>pro</sub> and related enzymes.

### 3.2. Specificity

Picornaviral 3C proteases generally cleave peptide sequences with a hydrophobic residue at P4, Gln at P1 and a small residue (Gly, Ser, Ala) at P1′ (Blom, Hansen, Blaa, & Brunak, 1996; Seipel et al., 1999). However, within that specificity framework there are some interesting variations. Although HRV and PV 3C<sub>pro</sub> almost invariably cleave sequences with P1′-Gly, HAV 3C<sub>pro</sub> tolerates larger residues at this position and, in addition,
exhibits a preference for a small hydrophilic residue at P2 (Seipelt et al., 1999). FMDV 3C<sup>pro</sup> shows similar variability at P1' but is unusual in also being able to cleave sequences with P1-Gln or P1-Glu at broadly similar rates (Birtley et al., 2005); this reflects the fact that in the viral polyprotein, there are five 3C<sup>pro</sup> cleavage junctions with P1-Gln and five with P1-Glu. Experimental analyses confirm the importance of P4, P1 and P1’ positions in peptides cleaved by FMDV 3C<sup>pro</sup> but also suggests important roles for P2 and P4’ positions (Birtley et al., 2005).

To analyse the cleavage specificity further, we aligned the FMDV 3C<sup>pro</sup> cleavage sequences recently reported for over 100 strains of the virus (Carrillo et al., 2005) and grouped them according to the identity of the P1 residue (Gln or Glu) (Fig. 2c). This reveals some previously undetected correlations between different positions in the sequences recognised by FMDV 3C<sup>pro</sup>. Thus, sequences with P1-Gln typically also have P2-Lys and a largely hydrophobic residue at P1’ (Leu, Ile, Thr). In contrast, in sequences with P1-Glu the preference for P2-Lys is reduced but there is strong selectivity for a small amino acid (Gly or Ser) at P1’. This suggests that interactions between different subsites in the peptide binding cleft of FMDV 3C<sup>pro</sup> may be important for specificity and, in particular, that recognition of the P1 residue is influenced by the P2 and P1’ positions. The sequence specificity identified for junctions with P1-Glu (Fig. 2c) helps to explain why the sequence VRAE/VQ in eIF4AII is cleaved by FMDV 3C<sup>pro</sup>, but the closely related sequence in eIF4AI (VRNE/MQ) is not (Li, Ross-Smith, Proud, & Belsham, 2001).

Further structural studies should help to elucidate the details of 3C<sup>pro</sup> specificity since, to date, there are no co-crystal structures of picornaviral 3C proteases with bound peptide and only a handful of reports of complexes containing peptide-like inhibitors (Bergmann et al., 1999; Dragovich et al., 2002; Matthews et al., 1999). At present we do not even properly understand the structural basis for selectivity of P1-Gln in most 3C proteases, since the residues that apparently specify P1-Gln in PV, HRV and HAV 3C<sup>pro</sup> are also conserved in FMDV 3C<sup>pro</sup> (Birtley et al., 2005), which can accommodate Gln or Glu at this position.

### 3.3. Cis/trans cleavage

Does FMDV 3C<sup>pro</sup> excise itself from the polyprotein precursor by cis or trans cleavages? Structural studies of the mature 3C proteases from other picornaviruses suggest that cis-cleavage at the N-terminus of 3C<sup>pro</sup> occurs more readily than cleavage at the C-terminus since it requires less distortion of the precursor to position the N-terminal cleavage junction in the active site (Bergmann et al., 1997; Khan, Khazanovich-Bernstein, Bergmann, & James, 1999; Matthews et al., 1994). This hypothesis is consistent with the requirement for a functional 3CD precursor (Leong et al., 2002). FMDV 3C<sup>pro</sup> may well behave similarly although it should be noted that it has a longer C-terminus (by 7–13 residues) than other picornavirus 3C proteases (Birtley et al., 2005), which may allow for enhanced manoeuvrability of the C-terminal cleavage junction in the precursor. Again this is an area where more work is needed; for example, the crystal structure of a FMDV 3CD precursor (proteolytically inactivated by mutagenesis) may help to test models of polyprotein processing since it should reveal the location of the cleavage junction prior to proteolysis.

### 4. Biological functions

Although the primary role of FMDV 3C<sup>pro</sup> is processing of the viral polyprotein, the enzyme also cleaves host proteins within infected cells. It removes the N-terminal 20 amino acids from histone H3, probably leading to a down-regulation of transcription (Falk et al., 1990). A similar depression of transcription is achieved by PV 3CD<sup>pro</sup> by a different route: this precursor protease targets both the TATA-box binding protein and transcription factor IIC (Skern et al., 2002). It is not known if different picornaviruses target similar components of the transcription machinery. Most picornaviruses also inhibit host–cell translation initiation, primarily by cleavage of eIF4G. Although the viral leader protease (L<sup>pro</sup>) cleaves eIF4G to shut-off host cell protein synthesis in FMDV-infected cells, FMDV 3C<sup>pro</sup> has also been reported to cleave both eIF4G and eIF4A (Belsham, McInerney, & Ross-Smith, 2000; Li et al., 2001). However, this occurs at relatively late stages of viral infection and its contribution to viral replication has yet to be elucidated.

More recent work has highlighted a non-proteolytic role played by FMDV 3C<sup>pro</sup> in the replication of viral RNA. During the initiation of replication of positive strand RNA synthesis, the 3B protein (VPg) is uridylated to generate a VPg-pU-pU precursor that acts as a primer for positive strand synthesis on a negative strand RNA template. The formation of this precursor is critically dependent on the assembly of a protein–RNA complex involving a conserved cis-acting replication element (located within the 5'-untranslated region of the FMDV genome), the precursor 3CD and the 3D polymerase (Nayak, Goodfellow, & Belsham, 2005). Follow-up studies indicate that the formation of a functional 3B uridylylation complex in FMDV infected cells depends...
on the RNA binding activity of the 3C moiety of 3CD (Nayak et al., 2006). The RNA binding surface, as in other picornavirus 3C proteases, maps to a conserved basic patch on the dorsal surface of the enzyme, opposite to the active site (Fig. 2d). This patch is centred on a generally conserved $^{95}\text{KFRD}^{99}$ sequence but also contains other basic residues. FMDV 3C$^{\text{pro}}$ can functionally substitute for 3CD within in vitro uridylylation assays but at much lower efficiency. Clearly, the structure of the complex would help elucidate the mechanism in greater detail.

5. Possible applications—a viable drug target?

The central roles that FMDV 3C$^{\text{pro}}$ plays in polyprotein processing and RNA replication are undoubtedly linked to the observation that it is one of the most highly conserved proteins in the viral genome (Carrillo et al., 2005). Amino acid variation in 3C$^{\text{pro}}$ between different strains of FMDV is largely confined to the surface of the enzyme away from the active site, making this enzyme an attractive target for antiviral drug design (Birtley et al., 2005), not least because inhibitors – in marked contrast to vaccines – are likely to be active against different strains of the virus. The elucidation of the structure of FMDV 3C$^{\text{pro}}$ paves the way for structure-based drug design. However, this is clearly a difficult road since the structures of HRV and HAV 3C$^{\text{pro}}$ were determined over 10 years ago and have yet to yield commercially viable antiviral drugs for the associated diseases. In the case of FMDV, the enormous economic impact of the disease, coupled with the problems associated with vaccine use and the fact that drugs might be used prophylactically to control outbreaks, may yet stimulate further investment in this direction.

Acknowledgements

We thank Graham Belsham and Tim Skern for critical reading of the manuscript. SC and RJL acknowledge grant support from the BBSRC. NR is funded by a Marie Curie Host Fellowship for Early Stage Research Training.

References

Allaire, M., Chernia, M. M., Malcolm, B. A., & James, M. N. (1994). Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. Nature, 369(6475), 72–76.

Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J., & Hilgenfeld, R. (2002). Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alphahelical domain. EMBO J., 27(13), 3213–3224.

Barrett, A. J., & Rawlings, N. D. (2001). Evolutionary lines of cysteine peptidases. Biol. Chem., 382(5), 727–733.

Belsham, G. J., McInerney, G. M., & Ross-Smith, N. (2000). Foot-and-mouth disease virus 3C protease induces cleavage of translation initiation factors eIF4A and eIF4G within infected cells. J. Virol., 74(1), 272–280.

Bergmann, E. M., Cherney, M. M., McKendrick, J., Fromrorn, S., Luo, C., Malcolm, B. A., et al. (1999). Crystal structure of an inhibitor complex of the 3C proteinase from hepatitis A virus (HAV) and implications for the polyprotein processing in HAV. Virology, 265(1), 153–163.

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

Birtley, J. R., Knox, S. R., Jaulent, A. M., Brick, P., Leatherbarrow, R. J., & Curry, S. (2005). Crystal structure of foot-and-mouth disease virus 3C protease: New insights into catalytic mechanism and cleavage specificity. J. Biol. Chem., 280(12), 11520–11527.

Blom, N., Hansen, J., Blaa, D., & Brunak, S. (1996). Cleavage site analysis in picornaviral polyproteins: Discovering cellular targets by neural networks. Protein Sci., 5(11), 2203–2216.

Brenner, S. (1988). The molecular evolution of genes and proteins: A tale of two serines. Nature, 334(6182), 528–530.

Carrillo, C., Tulman, E. R., Delhon, G., Lu, Z., Carreno, A., Vagnozzi, A., et al. (2005). Comparative genomics of foot-and-mouth disease virus. J. Virol., 79(10), 6487–6504.

Dragovich, P. S., Prins, T. J., Zhou, R., Brown, E. L., Maldonado, F. C., Fuhrman, S. A., et al. (2002). Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 6. Structure–activity studies of orally bioavailable 2-pyridone-containing peptidomimetics. J. Med. Chem., 45(8), 1607–1623.

Falk, M. M., Grigera, P. R., Bergmann, I. E., Multhaup, G., & Beck, E. (1990). Foot-and-mouth disease virus protease 3C induces specific proteolytic cleavage of host cell histone H3. J. Virol., 64(2), 748–756.

Grubman, M. J., & Baxt, B. (2004). Foot-and-mouth disease. Clin. Microbiol. Rev., 17(2), 465–493.

Grubman, M. J., Zellner, M., Bablanian, G., Mason, P. W., & Piccone, M. E. (1995). Identification of the active-site residues of the 3C proteinase of foot-and-mouth disease virus. Virology, 123(2), 581–589.

Hedstrom, L. (2002). Serine protease mechanism and specificity. Chem. Rev., 102(12), 4501–4524.

Kean, K. M., Teterina, N. L., Marc, D., & Girard, M. (1991). Analysis of putative active site residues of the poliovirus 3C protease. Virology, 175(2), 609–619.

Khan, A. R., Khazanovich-Bernstein, N., Bergmann, E. M., & James, M. N. (1999). Structural aspects of activation pathways of aspartic proteasezymogens and viral 3C protease precursors. Proc. Natl. Acad. Sci. U.S.A., 96(20), 10968–10975.

Leong, L. E.-C., Cornell, C. T., & Semler, B. L. (2002). Processing determinants and functions of cleavage products of picornavirus proteins. In B. L. Semler, & E. Wimmer (Eds.), Molecular Biology of Picornaviruses (pp. 187–197). Washington, DC: ASM Press.

Li, W., Ross-Smith, N., Proud, C. G., & Belsham, G. J. (2001). Cleavage of translation initiation factor 4AI (eIF4AI) but not eIF4G by foot-and-mouth disease virus 3C protease: Identification of the eIF4AI cleavage site. FEBS Lett., 507(1), 1–5.

Mason, P. W., Grubman, M. J., & Baxt, B. (2003). Molecular basis of pathogenesis of FMDV. Virus Res., 97(1), 9–32.
Matthews, D. A., Dragovich, P. S., Webber, S. E., Fuhrman, S. A., Patick, A. K., Zalman, L. S., et al. (1999). Structure-assisted design of mechanism-based irreversible inhibitors of human rhinovirus 3C protease with potent antiviral activity against multiple rhinovirus serotypes. *Proc. Natl. Acad. Sci. U.S.A.*, 96(20), 11000–11007.

Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W., et al. (1994). Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. *Cell*, 77(5), 761–771.

Mosimann, S. C., Cherney, M. M., Sia, S., Plotch, S., & James, M. N. (1997). Refined X-ray crystallographic structure of the poliovirus 3C gene product. *J. Mol. Biol.*, 273(5), 1032–1047.

Nayak, A., Goodfellow, I. G., & Belsham, G. J. (2005). Factors required for the uridylylation of the foot-and-mouth disease virus 3B1, 3B2, and 3B3 peptides by the RNA-dependent RNA polymerase (3Dpol) in vitro. *J. Virol.*, 79(12), 7698–7706.

Nayak, A., Goodfellow, I. G., Woolaway, K. E., Birtley, J. R., Curry, S., & Belsham, G. J. (2006). Role of RNA structure and RNA binding activity of the foot-and-mouth disease virus 3C protein in VPg uridylylation and virus replication. *J. Virol.*, 80(19), 9865–9875.

Phan, J., Zdanov, A., Evdokimov, A. G., Tropea, J. E., Peters, H. K., III, Kapust, R. B., et al. (2002). Structural basis for the substrate specificity of tobacco etch virus protease. *J. Biol. Chem.*, 277(52), 50564–50572.

Sarkany, Z., & Polgar, L. (2003). The unusual catalytic triad of poliovirus protease 3C. *Biochemistry*, 42(2), 516–522.

Sarkany, Z., Szeltner, Z., & Polgar, L. (2001). Thiolate-imidazolium ion pair is not an obligatory catalytic entity of cysteine peptidases: The active site of picornain 3C. *Biochemistry*, 40(35), 10601–10606.

Seipel, J., Guarne, A., Bergmann, E., James, M., Sommergruber, W., Fita, I., et al. (1999). The structures of picornaviral proteases. *Virus Res.*, 62(2), 159–168.

Skern, T., Hampoelz, B., Guarne, A., Fita, I., Bergmann, E., Petersen, J., et al. (2002). Structure and function of picornavirus proteases. In B. L. Semler, & E. Wimmer (Eds.), *Molecular Biology of Picornaviruses* (pp. 199–212). Washington, DC: ASM Press.

Sweeney, T. R., Roque-Rosell, N., Birtley, J. R., Leatherbarrow, R. J., & Curry, S. (2006). Unpublished results.

Yin, J., Bergmann, E. M., Cherney, M. M., Lall, M. S., Jain, R. P., Vederas, J. C., et al. (2005). Dual modes of modification of hepatitis A virus 3C protease by a serine-derived beta-lactone: Selective crystallization and formation of a functional catalytic triad in the active site. *J. Mol. Biol.*, 354(4), 854–871.

Zeitler, C. E., Estes, M. K., & Prasad, B. V. (2006). X-ray crystallographic structure of the Norwalk virus protease at 1.5 ˚A resolution. *J. Virol.*, 50(10), 5050–5058.