Characterization of a protein from *Rice tungro spherical virus* with serine proteinase-like activity

Vera Thole¹ and Roger Hull²

¹John Innes Centre, Department of Metabolic Biology and Department of Disease and Stress Biology, Norwich Research Park, Norwich NR4 7UH, UK

The RNA genome of *Rice tungro spherical virus* (RTSV) is predicted to be expressed as a large polyprotein precursor (Shen et al., Virology 193, 621–630, 1993). The polyprotein is processed by at least one virus-encoded protease located adjacent to the C-terminal putative RNA polymerase which shows sequence similarity to viral serine-like proteases. The catalytic activity of this protease was explored using *in vitro* transcription/translation systems. Besides acting *in cis*, the protease had activity *in trans* on precursors containing regions of the 3′ half of the polyprotein but did not process a substrate consisting of a precursor of the coat proteins. The substitution mutation of Asp²⁷³⁵ of the RTSV polyprotein had no effect on proteolysis; however, His²⁶⁸⁰, Glu²⁷¹⁷, Cys²⁸¹¹ and His²⁸³⁰ proved to be essential for catalytic activity and could constitute the catalytic centre and/or substrate-binding pocket of the RTSV 3C-like protease.

Rice tungro spherical virus (RTSV) comprises a positive single-stranded RNA genome of about 12 kb contained within a capsid composed of three protein species. Analysis of the primary structure of RTSV RNA reveals one large open reading frame (ORF) potenially encoding a polyprotein of about 390 kDa and two small ORFs at the 3′ genomic end (Shen et al., 1993; Thole & Hull, 1996; Fig. 1A). The genetic organization of the large ORF resembles that of picornaviruses. Animal and plant picornaviruses initially express their genetic information as a polyprotein precursor that is processed by one or more virus-encoded proteases into gene products (see Hull, 2001). In the RTSV polyprotein, a 35 kDa protease has been identified adjacent to the putative RNA-dependent RNA polymerase (Thole & Hull, 1998).

Comparative sequence analyses have suggested that the RTSV 35 kDa protease is closely related to 3C (-like) proteases of members of the family *Picornaviridae* (Koonin & Dolja, 1993; Ryan & Flint, 1997; Shen et al., 1993; Spall et al., 1997) which share many characteristics with cellular chymotrypsin-like serine proteases (for reviews, see Dougherty & Semler, 1993; Ryan & Flint, 1997). Cellular serine proteases possess a catalytic triad of amino acids comprising His, Asp and Ser, which are located in similar overall structures. The nucleophilic Ser is responsible for cleavage and is often replaced by a functionally and spatially equivalent Cys in viral trypsin-like proteases (Dessens & Lomonossoff, 1991; Dougherty et al., 1989; Lawson & Semler, 1991; Margis & Pinck, 1992). The Asp of the active-site residues can be substituted by an equivalent Glu (Allaire et al., 1994; Dessens & Lomonossoff, 1991; Hämmerle et al., 1991; Matthews et al., 1994). Substrate specificity is determined by the substrate-binding pocket which may involve either several amino acids or only one. For serine (-like) proteinases this pocket is thought and, in part, experimentally and structurally confirmed to involve, several predominantly hydrophobic amino acids near the active Ser/Cys. A highly conserved His downstream of the nucleophilic centre is regarded as part of the substrate-binding pocket in some 3C-like proteases (Allaire et al., 1994; Bazan & Fletterick, 1988; Blair et al., 1996; Gorbalenya et al., 1989; Matthews et al., 1994).

Picornaviral 3C proteases display a high degree of substrate specificity (Nicklin et al., 1986; Pallai et al., 1989) and catalyse the majority of processing events, including the cleavage of the coat proteins (CPs). The CP polyprotein precursor is processed in *trans* either by a mature 3C protease (Clarke & Sangar, 1988; Harmon et al., 1992; Jia et al., 1991; Vakharia et al., 1987) or by a 3C precursor (Jore et al., 1988; Ypma-Wong et al., 1988).

In this paper, we report on some of the *trans*-activity properties of the RTSV 3C-like protease and on the mutational studies of conserved amino acids considered to define the active site and/or the substrate-binding pocket of the protease.

Previously, we demonstrated a proteolytic activity in the C-terminal part of the RTSV polyprotein by analysing the primary translation products of pBat1534 [91 kDa; nucleotides (nt) 7055–9498; Figs 1C and 2F] and pBatE1 (58 kDa; nt 7920–9498; Fig. 1C) (Thole & Hull, 1998). To study RTSV polyprotein-processing further, deletion mutants containing C-terminal deletions in/of the predicted RTSV 3C-like
Fig. 1. Illustration of the RTSV genome and constructs used in this work. (A) Genome organization of RTSV showing the large polyprotein-encoding ORF and two small ORFs (adapted from Shen et al., 1993). The poly(A) tract is indicated at the 3′ end of the genome. The regions of the three coat proteins (CP 1, 2 and 3), the 35 kDa protease (Pro) showing sequence similarity with 3C-like proteases, domains for an NTP-binding protein (NTP) and an RNA-dependent RNA polymerase (Pol) are indicated in the polyprotein. Question marks denote putative protein products of unknown function. (B) Schematic diagram showing the region of the RTSV genome in construct pBatCP (single line) used for expressing CPs in relation to the RTSV genome and details of the CPs (boxes). Amino acid pairs probably forming processing sites (Shen et al., 1993; Zhang et al., 1993; Thole & Hull, 1998) are indicated in standard single-letter code above the genome together with the amino acid numbering (B–D). The nucleotide numbering of (putative) cleavage sites is shown below the genome. The amino acid and nucleotide numbering are based on the full-length RTSV sequence of Shen et al. (1993) and the RTSV nucleotide sequence used in our analyses is described in Thole & Hull (1996). (C) Schematic representation of the wild-type protease constructs pBat1534 and pBatE1, the cleavage-site mutant pBatQ2526P (altered at the N-terminal cleavage site Gln2526/Asp of the 3C-like protease) and the deletion mutants pBat15, pBat11 and pBat9 relative to the RTSV polyprotein. (D) Schematic localization of constructs with mutations of putative catalytic sites (pBatH2680G, pBatE2717Q, pBatD2735E, pBatC2811A and pBatH2830E) in relation to the RTSV genome together with the indication whether cleavage occurred or not.
protease domain (pBatΔ9, pBatΔ11 and pBatΔ15, Fig. 1C) were constructed by exonuclease III and restriction enzyme digestions of pBat1534. RTSV polypeptides were expressed in coupled in vitro transcription/translation systems using either wheat germ extract or rabbit reticulocyte lysate as translation systems and in the absence/presence of the reducing agent dithiothreitol as described by Thole & Hull (1998). Polypeptide synthesis was analysed under different conditions as proteases can be inactive in a certain translation system and/or require reducing agents for expression and (full) catalytic activity (Mavankal & Rhoads, 1991; Pelham, 1979; Shih et al., 1987; Verchot et al., 1991, 1992). Independently of the experimental conditions, the transcription/translation of pBatΔ9 (nt 7055–8136), pBatΔ11 (nt 7055–8520) and pBatΔ15 (nt 7055–9001) for 0.5–24 h yielded stable polypeptides of 41, 55 and 73 kDa, respectively (Fig. 2A–C and data not shown). Although it covers most of the protease gene (nt 8090–9068), the region spanning nt 7055–9001 did not yield active protease, which points to the key role of the amino acids encoded by the stretch spanning nt 9001–9068 in determining the protease function (Fig. 1A).

Based on the lack of a detectable proteolytic activity in the deletion mutant polypeptides, these polypeptides were used in trans-cleavage assays as substrates for the protease polypeptide precursors of pBatE1 and pBat1534. Trans-assays were performed with substrates translated in the presence of radioactively labelled methionine and enzyme sources synthesized using unlabelled methionine. The protease forms derived from pBat1534 and pBatE1 partially cleaved in trans the polypeptides synthesized from pBatΔ9, pBatΔ11 and pBatΔ15 to a similar extent and gave virtually identical product patterns (Fig. 2A–C). The efficiency of trans-cleavage was increased at a substrate:protease ratio of 1:10 compared to the ratios 1:1, 1:3 or 1:5, which is consistent with the dilution sensitivity of a bimolecular reaction. Comparing the products arising during trans-cleavage of the largest pBat1534 deletion mutant, pBatΔ15, with the cleavage pattern observed following self-processing of pBat1534 (Fig. 2C, F), it seems that in trans only certain sites were cleaved and with lower efficiency.

Besides the wild-type protease forms, the cis-processing-deficient cleavage-site mutant pBatQ2526P (nt 7055–9498; Fig. 1C; Thole & Hull, 1998) was used as an enzyme for cleaving the deletion mutant polypeptides. Trans-processing assays resulted in a cleavage pattern identical to that of the wild-type protease forms; however, two products of about 21 and 29 kDa were processed to a higher yield by pBatQ2526P (Fig. 2A–C). Thus, it appeared that the protease in the large non-self-processing polypeptide form cleaved several sites more efficiently in trans than the wild-type protease forms. Differential proteolytic activities of precursor and mature protease forms were described for several proteases (Dessens & Lomonossoff, 1992; Hall & Palmenberg, 1996; Jore et al., 1988; Margis et al., 1994; Ypma-Wong et al., 1988). In trans- assays where the pBatQ2526P polypeptide served as substrate for the wild-type protease forms pBatE1 and pBat1534 (Fig. 2D), processing at some sites only was demonstrated. This was in comparison with the self-processing profile of pBat1534 (Fig. 2F) and the product pattern was almost identical to that observed by trans-cleavage of the substrate pBatΔ15 (Fig. 2C) indicating that particular site(s) are (preferentially) cleaved in cis.

The nature of the RTSV CP cleavage sites led to the prediction that they might be processed by a 3C-like protease (Shen et al., 1993; Zhang et al., 1993). To explore this, the trans-processing activity of the RTSV 3C-type protease was analysed on a polypeptide substrate consisting of the three RTSV CPs without the predicted C terminus of CP 3 (nt 2117–4388; pBatCP; Fig. 1B). pBatCP was constructed by removing the CP-encoding fragment from pB3ea6 (Zhang et al., 1993) and joining it in-frame with the translational start consensus sequence. As a prerequisite for the trans-assays, it was demonstrated that the substrate pBatCP (83 kDa) had no catalytic activity in either wheat germ (data not shown) or rabbit reticulocyte translation systems (Fig. 2E), which also suggests that an (entire) protease is absent in the RTSV polyprotein region encoded by nt 2117–4388. In the trans-assays, the CP polyprotein precursor was co-incubated with the wild-type (pBatE1 and pBat1534) or mutant (pBatQ2526P) protease forms at different substrate:protease ratios (1:1, 1:3, 1:5 and 1:10) for 3 to 24 h in the presence/absence of a reducing agent in both animal and plant cell-free translation systems (Fig. 2E and data not shown). The proteolytic activity of the enzyme sources used was simultaneously verified by trans-cleavage of the pBatΔ15 polyprotein (data not shown). In all our trans-assays, processing of the RTSV CP polyprotein was not detectable which might be because of the following reasons. (1) An inhibitor(s) prevented the release of the CPs. (2) The protease and/or substrate required a specific precursor form for cleavage, e.g. a protease–polymerase intermediate (Jore et al., 1988; Ypma-Wong et al., 1988). This could not be investigated due to the instability of pBat clones containing the RTSV 3C-like protease and polymerase domain (Thole & Hull, 1998). (3) The protease may require accessory protein(s)/factor(s) that modulate cleavage at particular sites (Blair et al., 1993; Chambers et al., 1991; Failla et al., 1994; Peters et al., 1992). (4) Alternatively, the RTSV CP polyprotein might be processed only/most efficiently in a cis-dependent fashion (Carrington et al., 1989; Jia et al., 1991; Palmenberg et al., 1992; Verchot et al., 1992) or by another protease.

To identify the putative active site and/or substrate-binding pocket of the protease residing in the C-terminal part of the RTSV polyprotein, mutational analyses of conserved amino acids were carried out. The conserved residues were selected based on computer analyses of cellular serine proteases and viral serine-like cysteine proteases (Bazan & Fletterick, 1988; Gorbalenya et al., 1989; Koonin & Dolja, 1993), the crystal structure of two 3C proteases (Allaire et al., 1994; Matthews et al., 1994) and mutational studies of viral
proteases (Cheah et al., 1990; Dessens & Lomonossoff, 1991; Dougherty et al., 1989; Hämmerle et al., 1991; Margis & Pinck, 1992). Substitution mutations were individually introduced into construct pBat1534 by inserting PCR fragments containing single amino acid alterations as described by Thole & Hull (1998). The transversions for candidates of the active catalytic triad were His^{268} to Gly (5’ CATCCAGGTAACC TGCAGGCAT; pBatH268G) (reverse), Glu^{271} to Gln
Fig. 2. Trans-cleavage assays: in vitro processing of deletion mutant substrates (A–C), a cleavage site mutant substrate (D) and an RTSV CP precursor substrate (E) by wild-type and mutant enzyme constructs. The substrate and protease sources were synthesized individually in the rabbit reticulocyte cell-free translation system for 3 h and then RNase A (1 µg/10 µl) was added to the enzyme sources. The [35S]methionine-labelled deletion mutant substrates of pBatΔ9 (A), pBatΔ11 (B) and pBatΔ15 (C), the labelled CP polyprotein substrate of pBatCP (E) as well as the labelled cleavage site mutant substrate of pBatQ2526P (D) were added to each of the unlabelled enzyme sources, pBatE1, pBat1534 and pBatQ2526P, respectively, at a molar ratio of 1:10. The mixtures were incubated at 30 °C for 6 and 17–20 h and then analysed by 12% SDS–PAGE and autoradiography. The substrates of pBatΔ9 (A), pBatΔ11 (B), pBatΔ15 (C), pBatQ2526P (D) and pBatCP (E) alone are shown before the start of co-incubation with the enzymes (3 h), simultaneously with the 6 h co-incubation (9 h) and at the end of the analysis (20–23 h). In order to compare the trans-cleavage patterns (A–D) to the self-processing pattern of pBat1534, the processing of the pBat1534 polyprotein alone was monitored during 0.5 to 6 h (F). The precursor polyproteins of pBatΔ9 (41 kDa), pBatΔ11 (55 kDa), pBatΔ15 (73 kDa), pBatQ2526P (91 kDa), pBatCP (83 kDa) and pBat1534 (91 kDa) are marked by arrows and the molecular mass markers (in kDa) are indicated on the right of the autoradiographs. The (intermediate) processing products are identified by asterisks.

(5' CCCACACAACTGTCTGTTGAAATCC; pBatE2717Q) (reverse), Asp2735 to Glu (5' GCCACCCAGTCGAATTATTGGATTTATGC; pBatD2735E) (forward) and Cys2811 to Ala (5' ATGCCAGCTTTGCGAGGCTGTAT; pBatC2811A) (forward) (Fig. 1D). Further, the conserved His2830 which is thought to constitute part of the substrate-binding pocket was analysed by conversion to Glu (5’ TAATAGGAATGGAGCAGGTTGCGC; pBatH2830E) (forward) (Fig. 1D). The alterations of Glu2717, Cys2811 and His2830 which were analysed in a time-course from 0.5 to 20–24 h (Fig. 3A and data not shown) abolished each processing. The mutation at His2830 led to a severe limitation of processing, with a delayed and very partial cleavage after about 2–3 h of incubation (Fig. 3A). The substitution of Asp2735 did not have any detectable effect on proteolysis (Fig. 3A compared with control in Fig. 2F).

The mutant constructs were tested for their trans-cleavage properties by using the pBatΔ11-derived polypeptide as a substrate (Fig. 3B). Corresponding to the effects on their cis-proteolytic activities, the mutated proteases of pBatE2717Q, pBatC2811A and pBatH2830E did not act in trans on the deletion mutant polypeptide. Further, intermolecular activity of pBatH2680G was not detected in our assays. When Asp2735 was replaced by Glu the trans-activity was at wild-type level.

To examine whether the eliminated proteolytic activity was due to protein misfolding, the conserved amino acid mutants served as substrates for the wild-type protease pBat1534 in trans-assays. The mutant polyproteins were partially cleaved to several consistent products suggesting that the formation of an active enzyme was not prevented by (merely) altering the protein structure with the conversion of the amino acids (data not shown).

Our results suggest that the RTSV protease, located in the C-terminal half of the polyprotein, could belong to the group of serine-like cysteine proteinases and that it contains at its catalytic site the conserved residues Glu2717 and Cys2811. The potential active-site candidate Asp2735, analysed based on the prediction of Bazan & Fletterick (1988), does not constitute part of the catalytic centre. Glu2717 is consistent with the model of Gorbalenya et al. (1989) and mutational analyses of polioviral 3C proteases (Hämmerle et al., 1991; Kean et al., 1991). The almost-fully destroyed cis-activity and the diminished trans-activity of the substitution mutation at His2830 confirm its proposed role in an efficient catalysis but its role cannot be assessed with certainty as the third residue of the catalytic triad. Alterations of catalytic site residues can result in partial processing (Carter & Wells, 1988; Dessens & Lomonossoff, 1991; Dougherty et al., 1989; Snijder et al.,
Fig. 3. Self-processing of in vitro transcription/translation products produced by clones mutated in presumptive candidates for the active site and/or substrate-binding pocket of the RTSV 3C-type protease (A) and intermolecular assays of these mutant protease constructs (B). (A) The 35S-labelled animal cell-free translation products of pBatH2680G, pBatE2717Q, pBatD2735E, pBatC2811A, and pBatH2830E were analysed 0.5, 1, 2, and 3 h, 6 or 7 h and/or 20 h after start of the reaction by 12% SDS–PAGE and autoradiography. The polyprotein precursors of the pBats (each 91 kDa) are indicated by arrows and the position of the molecular mass markers (in kDa) is shown on the right. For pBatH2680G, two cleavage products are identified by asterisks. (B) The labelled substrate synthesized from pBatΔ11 was co-incubated with the pBatH2680G-, pBatE2717Q-, pBatC2811A-, pBatH2830E-, pBatD2735E-derived unlabelled mutant enzymes as well as the wild-type enzyme source of pBat1534 in a substrate:protease ratio of 1:10 after 3 h of coupled transcription/translation in rabbit reticulocyte lysate for each construct. Samples of the substrate–protease mixes were taken after 6 and 20 h of co-incubation at 30 °C and analysed by 12% SDS–PAGE followed by autoradiography. The pBatΔ11 substrate alone was monitored at 3 h (i.e. before the start of co-incubation with the enzymes), at 9 h (i.e. after 6 h of co-incubation) and at 23 h (after 20 h co-incubation). The polyprotein precursor pBatΔ11 (55 kDa) is marked by an arrow and molecular sizes (in kDa) are presented on the right. Cleavage products are identified by asterisks.

V. Thole and R. Hull

1996) and mutations can also have differential effects at particular cleavage sites (Kean et al., 1991, 1993). The substitution at His2830 of the RTSV polyprotein abolished processing, indicating its probable importance in substrate binding by analogy to RTSV His2830 equivalents in 3C-like proteases (Blair et al., 1996; Cheah et al., 1990; Hans & Sanfaçon, 1995; Ivanoff et al., 1986; Lawson & Semler, 1991; Snijder et al., 1996). This is in accord with the suggestion that when His is present in the binding pocket of a 3C-like proteinase, the proteinase seems to prefer Gln or Glu at the −1 position (Bazan & Fletterick, 1988; Gorbalenya et al., 1989; Allaire et al., 1994; Matthews et al., 1994).
We thank Dr S.-C. Lee for the stimulating discussions. The work was supported by the Rockefeller Foundation and the Overseas Development Administration Natural Resources Institute. RTSV was held under MAFF licence PHF 1185A/56 (110) and genetic manipulations were carried out under licence number PHF 1419A/982/34. JIC is grant-aided by the BBSRC.

References


