

# Activity of a purified His-tagged 3C-like proteinase from the coronavirus infectious bronchitis virus

K.W. Tibbles<sup>1,\*</sup>, D. Cavanagh<sup>2</sup>, T.D.K. Brown<sup>1</sup>

<sup>1</sup> Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

<sup>2</sup> Institute for Animal Health, Compton Laboratory, Newbury, Berks RG20 7NN, UK

Received 23 October 1998; accepted 12 January 1999

---

## Abstract

Previous studies *in vitro* of the processing of cloned polyprotein fragments from the coronavirus infectious bronchitis virus (IBV) large open reading frame (ORF1), confirmed the activity of a predicted 3C-like proteinase (3CLP) domain and suggested that the proteinase is released autocatalytically from the polyprotein in the form of a 35 kDa protein, 3CLpro, capable of further cleavages *in trans*. In order to identify such cleavages within the ORF1 polyprotein mediated by 3CLpro, the proteinase was expressed in bacteria, purified and used in *trans* cleavage assays with polyprotein fragments lacking the 3CLP domain as targets. The proteinase was expressed as a polyprotein fragment which was able to process during expression in bacterial cells, releasing mature 3CLpro. A histidine (His<sub>6</sub>) tag was introduced close to the C-terminus of the proteinase to aid purification. Processing demonstrated by the tagged proteinase was indistinguishable from that of the wild-type enzyme indicating that the site chosen for the tag was permissive. From these studies we were able to demonstrate *trans* cleavages consistent with the use of most of the previously predicted or identified sites within the open reading frame of gene 1. This tentatively completes the processing map for the ORF1 region with respect to 3CLpro. © 1999 Elsevier Science B.V. All rights reserved.

**Keywords:** Coronavirus; 3CLproteinase; His-tagged; Bacterial expression; *Trans* processing

---

## 1. Introduction

The coronavirus infectious bronchitis virus (IBV) encodes the major portion of its non-struc-

tural proteins within two large open reading frames, ORFs 1a and 1b, situated at the 5' end and occupying the major portion of the 27.6 kb single-stranded positive-sense RNA genome (Brown and Brierley, 1995). The ORFs overlap slightly with ORF1b accessed via programmed ribosomal frameshifting (Brierley et al., 1989) resulting in ORF1a and ORF1a/1b (a fusion) trans-

---

\* Corresponding author. Tel.: +44-1223-336917; fax: +44-1223-336926.

E-mail address: kwt@mole.bio.cam.ac.uk (K.W. Tibbles)

lation products. The large size of the ORFs suggests that the translation products are polyproteins that are processed to generate a set of mature non-structural proteins. In common with all other coronaviruses so far examined, IBV encodes a 3C-like proteinase domain which is located towards the C-terminus of the ORF1a polyprotein and approximately centrally in the ORF1a/1b fusion product. The 3CLP domain was first identified as a possible component of IBV polyproteins (Gorbalenya et al., 1989) due to sequence similarity to picornavirus 3C proteinases (reviewed in Ryan and Flint, 1997). In the same study, predictions were made as to likely target dipeptide cleavage sites for a 3C-like proteinase within the ORF1 polyproteins, the large number suggested implicating this proteinase in a pivotal role in the maturation of the polyproteins. Similar domains were subsequently identified in mouse hepatitis virus (MHV) (Lee et al., 1991) and human coronavirus (HCV 229E) (Herold et al., 1993). There are no cellular counterparts for the activities of the 3C-like proteinases, therefore they are essential to the successful maturation of their respective polyproteins and subsequent replication events. This, and the comparatively high degree of conservation (at functional level), makes the 3C-like proteinases attractive targets for antiviral intervention.

Given the essential role played by the 3C-like proteinase of the coronaviruses in maturation of the 'polymerase' polyproteins, there has been considerable interest in characterising the enzyme, with initial emphasis on defining the processing events involving the proteinase. Such studies are already at an advanced stage for both IBV (Liu et al., 1994, 1998) and HCV 229E ORF1b regions (Grotzinger et al., 1996; Heusipp et al., 1997a,b). 3C-like proteinase activities have been confirmed recently for three representatives of the coronavirus group, MHV (Lu et al., 1995), HCV 229E (Ziebuhr et al., 1995) and IBV (Tibbles et al., 1996) and the protein has been identified *in vivo* during viral infection in the case of HCV (Ziebuhr et al., 1995) and MHV (Lu et al., 1996). In addition, the proteinases from these two viruses have been purified and partially characterised (Seybert et al., 1997; Ziebuhr et al., 1997). The

identification of the IBV 3CLpro was first realised *in vitro* when a cloned fragment of the polyprotein was expressed in reticulocyte lysate (Tibbles et al., 1996). In the presence of membranes, upon which activity was vitally dependent, processing of the region containing the protease domain was observed with the release of a 35 kDa protein (p35). This product was capable of *trans* processing, but was inefficient and of limited use for mapping studies. We undertook the identification of the remaining IBV 3CLpro cleavage targets within the ORF1 polyprotein using bacterial expression to obtain significant quantities of 3CLpro.

## 2. Expression and purification of 3CLproH6

The bacterial expression system described utilises the minimum processing unit identified in our previous studies which established that, in addition to the 3CLP domain, MP2 protein sequence between the Q/S<sub>4</sub> cleavage target and a point delimited by an *Nco*I restriction site (ntd position 10118) was that minimally required for processing. A *Bam*HI fragment released from pKT205 (see below) was introduced into *Bam*HI digested pET3xa (Novagen) to yield pET205 (Fig. 1). This construct would allow expression to be monitored through the synthesis of the strongly expressed T7 gene 10 protein, which composes the major insoluble fraction component upon extraction. Expression beyond the *Nco*I site was restricted by digestion of pET205 with *Nco*I followed by end repair and religation. This manipulation introduces a frame-shift in the sequence which results in a translation stop signal being encountered. This plasmid was then used to create an internal deletion between the vector initiator ATG triplet and the IBV sequence by digestion with *Nhe*I and *Bam*HI. The newly generated plasmid, pETAN205ΔC<sup>Nco</sup>, was modified to include a His-6 tag approximately 35 amino-acid residues from the C-terminus of 3CLpro to aid in purification. A mutation which replaced the amino-acid sequence VKNSQW (single letter code) with six histidine residues near the C-terminus of the 3CLP domain in pKT205 (ntds 9679-

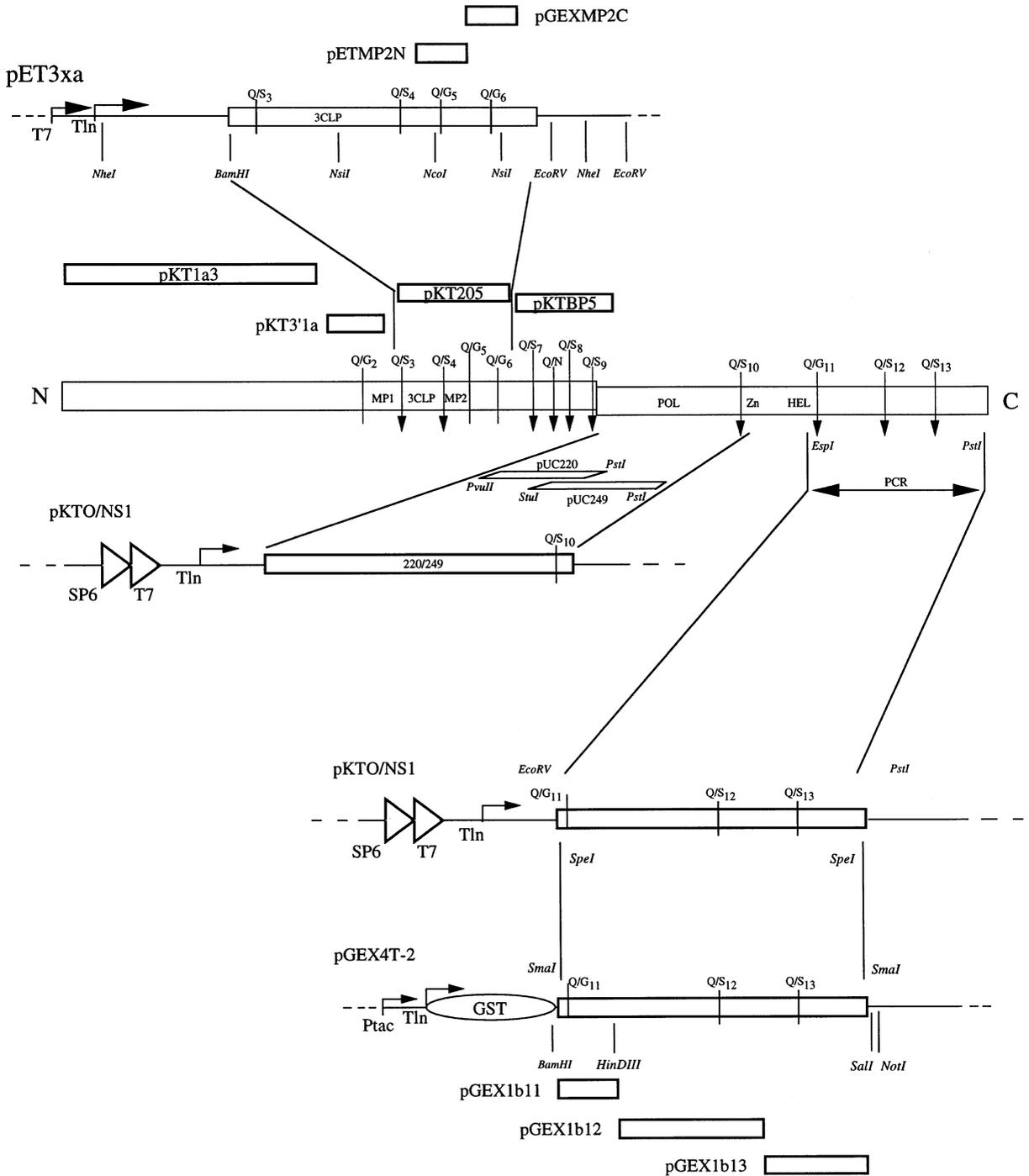


Fig. 1. Schematic representation of the IBV ORF1a/1b polyprotein with positions of predicted cleavage dipeptide targets shown as solid vertical bars, those with arrowheads have been experimentally confirmed. Predicted catalytic domains are also indicated (MP1/2; membrane proteins, 3CLP; 3C-like proteinase, POL; polymerase, Zn; zinc finger, HEL; helicase). The derivation of the eukaryotic and bacterial expression clones and the principal cloning vectors used are depicted. The principal restriction sites used in cloning manipulations (adjacent solid bars), promoters (T7, SP6 and Ptac) and translation (Tln) start sites for the plasmids are also shown.

9696) was made, giving rise to pKT205His. The positioning of this tag sequence was determined by consideration of the possible amino-acid sequences involved in cleavage at the Q/S<sub>4</sub> site and potential for interaction with the affinity matrix based on predicted hydrophilicity (Kyte and Doolittle, 1982) of the proteinase sequence. This modified sequence was transferred to the 205-bearing pET plasmid via replacement of a *Nsi*I fragment. The processing properties of the modified proteinase were tested extensively, using a variety of systems, and were found to be indistinguishable from wild-type (data not shown). Therefore the appropriate pET plasmids (Studier et al., 1990) were introduced into *E.coli* BL21 pLysS cells (Studier, 1991). Cultures were grown for approximately 3 h until the OD<sub>600</sub> was between 0.6 and 0.9 and expression induced by the introduction of IPTG (at 100 µg/ml) into the growth medium. The cells were harvested after a further 3 h. Cells were lysed, clarified and the supernatant applied to a chelating agarose column charged with cobalt chloride. The use of cobalt instead of the usual nickel resulted in lower binding of background material (data not shown). Alternatively, cultures were induced after 1.5–2 h

(OD<sub>600</sub> 0.3 approx.) and rifampicin was added (to 200 µg/ml) after a further hour and incubation continued for 3 h. This resulted in a substantial inhibition of cellular protein expression without affecting ongoing T7-dependent expression. Although this resulted in a lower overall protein yield, there was a higher proportion of 3CLpro and background binding for the metal chelate column selection was significantly reduced. Typical results are shown in Fig. 2 where the appearance of a novel 35 kDa protein was observed after induction (compare lanes 1 and 2, panel A). Clarification of the crude lysate (lane 3) by centrifugation revealed that the major portion of the protein of interest (i.e. 35 kDa) remained in the supernatant (lane 5, compare to lane 4, pellet fraction). Comparison of the total applied protein and the unbound fraction (lanes 5 and 6 respectively) revealed that over 90% of the novel band seen after induction was retained on the column. This putative 3CLproH6 bound with relatively low affinity and was seen to elute at 60mM imidazole (lane 8), conditions normally associated with removing background binding. No observable protein remained to be eluted with 1M imidazole (lane 9). The background contaminating proteins

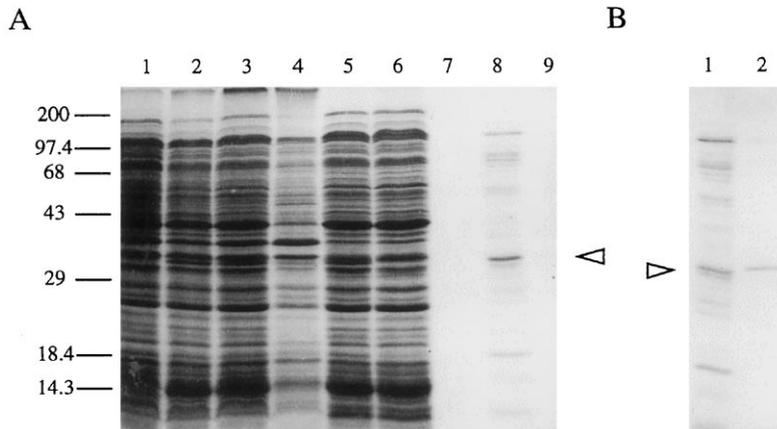


Fig. 2. Expression and purification of 3CLproH6. A. Samples are total lysates from noninduced (lane 1) and IPTG-induced (lane 2) cultures of *E.coli* BL21 cells transformed with pETΔN205ΔC<sup>Nco</sup>His or fractions made during purification. These are the total lysate (lane 3), pellet (lane 4) and clarified supernatant (lane 5) before application to the cobalt column, the unbound fraction (lane 6) and eluted fractions at 5 mM (lane 7), 60 mM (lane 8) and 1M (lane 9) imidazole. B. Comparison of the eluted proteins during normal induction (lane 1, equivalent to lane 8, panel A) and during induction in the presence of rifampicin (lane 2). All samples were normalised before separation on 17.5% polyacrylamide gels which were stained with Coomassie brilliant blue. The position of the induced 35 kDa product is indicated by open arrowheads.

could be significantly reduced by expression in the presence of rifampicin. While this resulted in a reduction in the total protein yield of some 80%, the proportion that was 3CLproH6 was greater and resulted in less background in the purified product (panel B, compare products under normal conditions, lane 1, with those obtained in the presence of rifampicin, lane 2, after single step purification on a cobalt column). Only fractions seen to contain the putative 3CLproH6 were found to cleave target polypeptide targets in a *trans* assay which supports the identity of the induced product as 3CLproH6 (data not shown). From comparison with known concentration protein standards (BSA) we estimated the yield to be approximately 2–4 mg/l of culture (0.5–1 mg/l in the presence of rifampicin), which compares favourably with that obtained by Ziebuhr et al., 1997.

### 3. Identification of possible 3CLpro target sites

Different regions of the ORF1a/1b polyprotein were expressed in reticulocyte lysate (TnT, Promega) or bacterial cell extract (S30, Promega) in the presence of [<sup>35</sup>S] methionine using the following expression plasmids (see Fig. 1). Construction of the pKT205 and pKTBP5 plasmids (encoding IBV sequence from nucleotides (ntd) 8693–10927 and 10752–12600 respectively) has been described elsewhere (Liu et al., 1994). These encode predicted cleavage sites Q/S<sub>3</sub> to Q/G<sub>6</sub> and Q/S<sub>7</sub> to Q/S<sub>9</sub>, respectively, within the C-terminal region of ORF1a. IBV sequence between ntd 16785 and 23060 (Boursnell et al., 1987), amplified by RT-PCR using a first-strand cDNA synthesis kit (Pharmacia) and *Pfu* polymerase (Stratagene), was cloned into pCRScript (Stratagene). This plasmid, pCRF4, was digested with *EspI* and *PstI* (following end repair of *EspI* site with DNA polymerase) and the fragment introduced into *EcoRV*–*PstI* digested pKT0/NS1 vector (Tibbles et al., 1995) to produce pKT1bC (encoding nucleotides 16841–20926). The polyprotein fragment encoded contains the three predicted 3CLpro cleavage sites Q/G<sub>11</sub>, Q/S<sub>12</sub> and Q/S<sub>13</sub> from the C-terminal region of ORF1b.

Cleavage dipeptides were mutated (to G/S or A/S) by oligonucleotide-directed mutagenesis as described previously (Tibbles et al., 1996). This region was transferred as a *SpeI* fragment to pGEX4T2 (Pharmacia) and deletions made to express portions containing each target dipeptide individually. The N-terminal half of the ORF1a region (up to nucleotide position 6500 approx.) is encoded by the plasmid pKT1a3, constructed as described previously (Liu et al., 1995). Analysis of the Q/G<sub>5</sub> site was made using a *KpnI* deletion of the pET205 plasmid, pET205Δ<sup>K<sub>pn</sub></sup>. A *HincII* fragment (ntd 10438–10902) bearing the Q/G<sub>6</sub> target from pKT205 was introduced into pGEX4T2 to yield pGEXMP2C (resulting in the production of a fusion with glutathione-S-transferase (GST), Smith and Johnson, 1988). The region encoding the MP1 region and Q/G<sub>2</sub> was derived from PCR amplification and a fragment (ntds 6919–8693) cloned upstream of the 205 sequence in pKT205/BP5.

Purified proteinase was added in varying amounts to the radiolabelled target fragments and incubated for up to 16 h at room temperature. Cleavage patterns were analysed following polyacrylamide gel electrophoresis and an example of these analyses is shown in Fig. 3. Complete digestion conditions for the fragment corresponding to the BP5 region of the polyprotein revealed a cleavage pattern that was consistent with the pattern expected from the cleavage predictions made by Gorbalenya et al., 1989. This target contained the predicted cleavage sites of Q/S<sub>7</sub>, Q/S<sub>8</sub> and Q/S<sub>9</sub>, all of which, along with an additional unpredicted Q/N site, have been experimentally confirmed (Liu et al., 1997; Ng and Liu, 1998). Cleavage within BP5 would be expected to yield processed products of approximate sizes of 6, 24, 10, 17 and 3 kDa (as they lie on the polyprotein) assuming cleavage at Q/S<sub>7</sub>, Q/N, Q/S<sub>8</sub> and Q/S<sub>9</sub>. In Fig. 3 the effect of processing by the addition of 3CLproh6 can be seen. The uncleaved target (lane 1) consisted of two species (of approximately 60 and 70 kDa) because the sequence contained the frame-shift signal, followed by approximately 300 nucleotides from ORF 1b. Only the two largest cleavage products were resolved in this experiment (lane 2) which correspond to cleavage

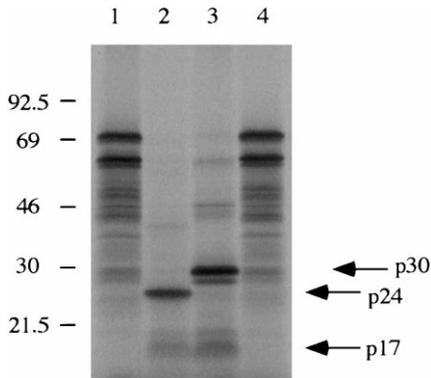


Fig. 3. *Trans* processing activity of 3CLproH6 against Q/S<sub>7</sub> target dipeptide. Samples of *in vitro* translations of the pK-TBP5-derived transcript in the presence of [<sup>35</sup>S] methionine were incubated in the absence (lane 1) or presence (lane 2) of 1 µg of purified 3CLproH6 for 16 h. Processing of [<sup>35</sup>S] methionine-labelled mutated BP5 target is shown in lane 3 and the control incubation in lane 4. Separation was by electrophoresis on 17.5% polyacrylamide gels and products detected by autoradiography. The positions of markers are indicated at left and processed products indicated by arrows to the right.

at Q/S<sub>7</sub> and Q/N for the largest product of 24 kDa and at Q/S<sub>8</sub> and Q/S<sub>9</sub> for the smaller of 17 kDa. The identities of these two products were

reinforced by comparing them with those produced after incorporation of [<sup>35</sup>S] cysteine. The relative intensity of the two bands is reversed reflecting the inverse ratio of the two amino acids in the 24 and 17 kDa products (data not shown). No cleavage was apparent in the absence of 3CLproH6 (data not shown). Cleavage at the predicted Q/S<sub>7</sub> site was tested by mutation of the cleavage dipeptide to an uncleavable form. Failure to cleave at the mutated Q/S<sub>7</sub> site should result in the extension of the N-terminus of the 24 kDa product by 6 kDa resulting in the generation of a new cleavage product of approximately 30 kDa. Comparison of the cleaved wild-type 24 kDa product (lane 2) and the product from the mutated polyprotein fragment (lane 3) revealed a decrease in the mobility of the more intensively labelled product (equating to approximately 30 kDa), consistent with failure to cleave at the mutated Q/S<sub>7</sub> site.

Polyprotein fragments constituting the remaining areas of the ORF1 region were synthesised and similarly analysed for cleavage by purified 3CLproH6. The results of all of these *trans* assay analyses are summarised in Table 1. We found that all so far identified target dipeptides are utilised by the 3CLproH6 enzyme *in vitro*. It was

Table 1  
Processing targets identified within the ORF1a/b polyprotein of IBV

Clone	Nucleotides	Expression system	Predicted site	<i>Trans</i> cleavage <i>in vitro</i>	Dipeptide mutation	Identification <i>in vivo</i>
pKT1a3	364–6495	RRL <sup>c</sup>	None	None	Irrelevant	No
pKT3'1a	6919–8448	RRL <sup>c</sup>	Q/G <sub>2</sub>	Not cleaved	Not done	No
pKT205	8693–10438	RRL <sup>c</sup>	Q/S <sub>3</sub>	Cleaved	Not cleaved	No
			Q/S <sub>4</sub>	Cleaved	Not cleaved	No
pETMP2N	10118–10438	S30 <sup>e</sup>	Q/G <sub>5</sub>	Not cleaved	Not done	No
pGEXMP2C	10438–10902	S30 <sup>e</sup>	Q/G <sub>6</sub>	Not cleaved	Not done	No
pKTBP5	10752–12600	RRL <sup>c</sup>	Q/S <sub>7</sub>	Cleaved	Not cleaved	Yes
			Q/N <sup>a</sup>	Cleaved <sup>b</sup>	Not done	Yes
			Q/S <sub>8</sub>	Cleaved <sup>b</sup>	Not done	Yes
			Q/S <sub>9</sub>	Cleaved <sup>b</sup>	Not done	Yes
pKT220/249	12625–15500	RRL <sup>c</sup>	Q/S <sub>10</sub>	Cleaved <sup>b</sup>	Not done	Yes
pGEX1b11	16854–17591	S30 <sup>e</sup>	Q/G <sub>11</sub>	Cleaved	Not cleaved	Yes
pGEX1b12	17591–19154	S30 <sup>e</sup>	Q/S <sub>12</sub>	Cleaved	Not cleaved	Yes
pGEX1b13	19154–20389	S30 <sup>e</sup>	Q/S <sub>13</sub>	Cleaved	Not cleaved	Yes

<sup>a</sup> Not previously predicted.

<sup>b</sup> Inferred from cleavage patterns only.

<sup>c</sup> RRL = rabbit reticulocyte lysate, S30 = bacterial lysate.

noted that under conditions expected to result in complete cleavage, not all target sites were digested to completion. These tended to be in ORF1b. Only three of the predicted sites fail to be cleaved, namely Q/G<sub>2</sub>, Q/G<sub>5</sub> and Q/G<sub>6</sub> and no novel target sites were detected. The first of the uncleaved sites (Q/G<sub>2</sub>) lies upstream of both the 3CLP and MP1 domains while the latter pair fall within the MP2 region. Either these sites are not utilised by 3CLpro or the conformation of the target fragments bearing them in our assays do not facilitate their recognition by the proteinase. However in the case of the Q/G<sub>6</sub>-bearing GST fusion, cleavage consistent with use of the thrombin site within the fusion junction region was observed (data not shown), which suggests that no gross conformational abnormality exists that might preclude proteolysis.

#### 4. Conclusions

We have used bacterial expression of viral protein, comprising the IBV ORF1a polyprotein fragment that gives rise to 3CLpro via autocatalytic processing in vitro (Tibbles et al., 1996), to obtain significant amounts of highly purified active proteinase for *trans* processing studies. Our rationale for modifying the starting pET plasmid containing the 205 sequence was as follows. Expression of hydrophobic proteins from picornaviruses with deleterious effects on bacterial cells has been reported (Lama and Carrasco, 1996). Therefore termination of expression at the *Nco*I site within MP2 would still allow processing, according to our in vitro data (Tibbles et al., 1996), while at the same time minimising the possible toxicity of the highly hydrophobic MP2 region. The gene 10 sequence was deleted in order to promote the production of a soluble polyprotein fragment, equivalent to that which is able to mature in vitro, and so ensure processing would proceed. Evidence suggests that in bacteria folding is post-translational (Netzer and Hartl, 1997) and so there was potential for the proteinase to become inactive due to the characteristic sequestration of pET fusion proteins into inclusion bodies. In the event this latter measure appeared

unnecessary as soluble processed 3CLpro was produced by the cells expressing fusion proteins (data not shown). This suggests that folding and processing were completed if not before completion of translation then before the sequestration into inclusion bodies. Thus by whatever means the proteinase matures in its usual eukaryotic context seems to be facilitated in the bacterial cell, though not necessarily by analogous interactions. The precise nature of the interaction between MP2 and cellular components (membranes) necessary for maturation of 3CLpro and its role in the mechanism of 3CLpro processing (maturation or release) remain to be determined. We do not know as yet whether the membrane requirement for maturation of 3CLpro also reflects a need for other associated cellular components such as chaperone functions. The fact that processing is able to proceed in an heterologous environment offers a possible means of investigating this further.

In order to simplify purification of 3CLpro, we introduced a histidine (His<sub>6</sub>) tag which facilitates separation on a suitable metal chelation matrix. The tag could not be placed at the extreme N or C-termini of the protein as is usual since the cleavage dipeptide and surrounding sequences are crucial for correct processing (Ryan and Flint, 1997). We introduced the tag into the nearest region of predicted hydrophilicity to the C-terminal processing site Q/S4 so as to minimise interference with the sequences involved in processing and allow interaction of the exposed histidine residues with the purification matrix. The tagged proteinase (3CLproH6) appears to be indistinguishable from wild-type and is readily purified to high purity. However, the affinity between modified proteinase and matrix is low and similar to background proteins suggesting that the tag may not be fully accessible. Our observation seems to conflict with the interpretation of a mutational analysis of HCV 229E 3CLpro by Ziebuhr et al., 1997 which concluded that the C-terminus of 3CLpro was important to function and intolerant of mutation. The His<sub>6</sub>-tag mutation here (a substitution) was well tolerated and would therefore appear to indicate a so called 'permissive site' in the C-terminus of the protein.

We have shown by *trans* cleavage analysis that processing occurs at the majority of the predicted 3CLpro cleavage dipeptides within the IBV ORF1 polyprotein and that no further cleavages are apparent. Our data reinforce recently reported *in vivo* studies and also partially overcome a potential shortcoming of these approaches where some products may not be consistently detectable, due to poor reactivity with antisera or short half life for example. This analysis tentatively completes the processing map of the ORF1 region with respect to 3CLpro. There appear to be no target sites for 3CLpro in the N-terminal half of the ORF1a polyprotein (covered by the 1a3 clone) so that processing of this region and possibly regions upstream is probably achieved through the PLP domain (Liu et al., 1995), alone or in combination with other as yet unidentified proteinase activities. Those target dipeptides that were not cleaved in our assays are associated with the hydrophobic domains. Cleavage of the MP1 region downstream of 3CLpro may be carried out by the PLP domain or an alternative activity, although failure for 3CLpro to cleave Q/G<sub>2</sub> may reflect a shortcoming in the presentation of the targets. The data presented here also suggest that there is no processing between Q/S<sub>3</sub> (the C-terminus of the 3CLP domain) and Q/S<sub>7</sub> so that the MP2 product containing uncleaved Q/G<sub>5</sub> and Q/G<sub>6</sub> dipeptides would be expected to be approximately 42 kDa. Notably these three hydrophobic region targets are not represented in either MHV or HCV 229E, although, along with IBV itself, these viruses bear a potential target at a position equivalent to just downstream of the IBV Q/G<sub>6</sub> site. These possibilities and the assembly of processed products into polymerase complexes may be addressed using specific antisera raised to amino or carboxyl terminal regions of processing products indicated by the present studies.

### Acknowledgements

We thank Ian Brierley for critical reading of the manuscript. This work was supported by the Wellcome Trust, the Biotechnology and Biological Sciences Research Council and the Ministry of Agriculture, Fisheries and Food, UK.

### References

- Bournsnell, M.E., Brown, T.D., Foulds, I.J., Green, P.F., Tomley, F.M., Binns, M.M., 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* 68, 57–77.
- Brierley, I., Digard, P., Inglis, S.C., 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* 57 (4), 537–547.
- Brown, T.D.K., Brierley, I., 1995. The coronavirus non-structural proteins. In: Siddell, S.G. (Ed.), *The Coronaviridae*. Plenum Press, New York, pp. 191–217.
- Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P., Blinov, V.M., 1989. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids. Res.* 17 (12), 4847–4861.
- Grotzinger, C., Heusipp, G., Ziebuhr, J., Harms, U., Suss, J., Siddell, S.G., 1996. Characterization of a 105-kDa polypeptide encoded in gene 1 of the human coronavirus HCV 229E. *Virology* 222 (1), 227–235.
- Herold, J., Raabe, T., Schelle Prinz, B., Siddell, S.G., 1993. Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. *Virology* 195 (2), 680–691.
- Heusipp, G., Grotzinger, C., Herold, J., Siddell, S.G., Ziebuhr, J., 1997a. Identification and subcellular localization of a 41 kDa, polyprotein lab processing product in human coronavirus 229E-infected cells. *J. Gen. Virol.* 78, 2789–2794.
- Heusipp, G., Harms, U., Siddell, S.G., Ziebuhr, J., 1997b. Identification of an ATPase activity associated with a 71-kilodalton polypeptide encoded in gene 1 of the human coronavirus 229E. *J. Virol.* 71 (7), 5631–5634.
- Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157 (1), 105–132.
- Lama, J., Carrasco, L., 1996. Screening for membrane-permeabilizing mutants of the poliovirus protein 3AB. *J. Gen. Virol.* 77 (9), 2109–2119.
- Lee, H.J., Shieh, C.K., Gorbalenya, A.E., Koonin, E.V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., Lai, M.M., 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* 180 (2), 567–582.
- Liu, D.X., Brierley, I., Tibbles, K.W., Brown, T.D., 1994. A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. *J. Virol.* 68 (9), 5772–5780.
- Liu, D.X., Shen, S., Xu, H.Y., Wang, S.F., 1998. Proteolytic mapping of the coronavirus infectious bronchitis virus 1b polyprotein: Evidence for the presence of four cleavage sites of the 3C-like proteinase and identification of two novel cleavage products. *Virology* 246, 288–297.
- Liu, D.X., Tibbles, K.W., Cavanagh, D., Brown, T.D.K., Brierley, I., 1995. Identification, expression, and processing of an 87-kDa polypeptide encoded by ORF 1a of the

- coronavirus infectious bronchitis virus. *Virology* 208, 48–57.
- Liu, D.X., Xu, H.Y., Brown, T.D., 1997. Proteolytic processing of the coronavirus infectious bronchitis virus 1a polyprotein: identification of a 10-kilodalton polypeptide and determination of its cleavage sites. *J. Virol.* 71 (3), 1814–1820.
- Lu, X., Lu, Y., Denison, M.R., 1996. Intracellular and in vitro-translated 27-kDa proteins contain the 3C-like proteinase activity of the coronavirus MHV-A59. *Virology* 2, 375–382 issn: 0042-6822.
- Lu, Y., Lu, X., Denison, M.R., 1995. Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. *J. Virol.* 69 (6), 3554–3559.
- Netzer, W.J., Hartl, F.U., 1997. Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388 (6640), 343–349.
- Ng, L.F., Liu, D.X., 1998. Identification of a 24-kDa polypeptide processed from the coronavirus infectious bronchitis virus 1a polyprotein by the 3C-like proteinase and determination of its cleavage sites. *Virology* 243 (2), 388–395.
- Ryan, M.D., Flint, M., 1997. Virus-encoded proteinases of the picornavirus super-group. *J. Gen. Virol.* 78 (4), 699–723.
- Seybert, A., Ziebuhr, J., Siddell, S.G., 1997. Expression and characterization of a recombinant murine coronavirus 3C-like proteinase. *J. Gen. Virol.* 78 (1), 71–75.
- Smith, D.B., Johnson, K.S., 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67 (1), 31–40.
- Studier, F.W., 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* 219 (1), 37–44.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., Dubendorff, J.W., 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 185, 60–89.
- Tibbles, K.W., Brierley, I., Cavanagh, D., Brown, T.D., 1996. Characterization in vitro of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus. *J. Virol.* 70 (3), 1923–1930.
- Tibbles, K.W., Brierley, I., Cavanagh, D., Brown, T.D.K., 1995. A region of the infectious bronchitis virus 1a polyprotein encoding the 3C-like protease domain is subject to rapid turnover when expressed in rabbit reticulocyte lysate. *J. Gen. Virol.* 76, 3059–3070.
- Ziebuhr, J., Herold, J., Siddell, S.G., 1995. Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. *J. Virol.* 69 (7), 4331–4338.
- Ziebuhr, J., Heusipp, G., Siddell, S.G., 1997. Biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase. *J. Virol.* 71 (5), 3992–3997.