Specificity of the polioviral proteinase 3C towards genetically engineered cleavage sites in the viral capsid

Caroline Mirzayan, Richard Ingraham and Eckard Wimmer

Department of Microbiology, State University of New York, Stony Brook, New York 11794 and Research and Development Center, Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, Connecticut 06877, U.S.A.

In a study of the cleavage specificity of poliovirus proteinase 3C<sup>pro</sup>, two mutant polioviruses were constructed to include putative 3C<sup>pro</sup> cleavage sites in the BC loop of VP1. The BC loop of VP1 in the wild-type virus is the neutralization antigenic site IA, consisting of a continuous chain of nine amino acids (ASTTNNKDKL). The first mutant, WI-ID-BC1, has four altered amino acids in the BC loop (ASTQGPGKL); the second mutant, WI-ID-BC2, has an insertion of nine amino acids in the BC loop (ASTGTAKVQGPGNKDKL). WI-ID-BC1 and WI-ID-BC2 were viable, grew to high titre and produced plaques of normal size. WI-ID-BC1 virions were resistant to proteolytic cleavage of the BC loop in vivo as well as upon incubation with a large excess of 3C<sup>pro</sup> in vitro, although a synthetic decapeptide (PASTQGPGKL) containing the amino acids of the BC loop in WI-ID-BC1 was cleaved by 3C<sup>pro</sup>. In contrast, WI-1D-BC2 yielded virus the VP1 of which was cleaved partially in vivo and completely when incubated with 3C<sup>pro</sup> in vitro. Our results showed that an insertion of nine amino acids into the antigenic loop of poliovirus, representing a synthetic 3C<sup>pro</sup> cleavage site, renders the loop susceptible to cleavage by proteinase 3C<sup>pro</sup>, but that this cleavage is restricted if the loop is the length of that in the native virion. This result implies that, in this case, structural restrictions override sequence determinants for cleavage of the BC loop by 3C<sup>pro</sup>.

Proteolytic processing plays a pivotal role in the regulation of gene expression of many positive-strand RNA viruses. The strategy employed by poliovirus, the best studied member of the Picornaviridae family, is to synthesize a single translation product encoding all the viral proteins and to regulate the expression of each protein by controlled proteolysis.

A genomic map of the poliovirus polyprotein, showing the deduced cleavage sites recognized by viral proteinases, is depicted in Fig. 1. Proteolysis occurs in three stages. The primary cleavage is carried out by proteinase 2A<sup>pro</sup> at a Tyr-Gly site, leading to the separation of the capsid proteins (P1) from the non-capsid proteins (P2 and P3) (Toyoda <i>et al.</i>, 1986). The secondary cleavages are catalysed by 3C<sup>pro</sup> and its precursor, 3CD<sup>pro</sup>, at Gln-Gly sites. Whereas 3C<sup>pro</sup> processes the non-capsid proteins, 3CD<sup>pro</sup> processes the capsid proteins (Jore <i>et al.</i>, 1988; Ypma-Wong <i>et al.</i>, 1988b). Lastly, the tertiary cleavage occurs during the morphogenesis of the virion cleaving VP0 at an Asn-Ser pair to produce VP4 and VP2 (Arnold <i>et al.</i>, 1987; Larsen <i>et al.</i>, 1982; Harber <i>et al.</i>, 1991). Although it is not yet clear what catalyses the cleavage of VP0, the involvement of either 2A<sup>pro</sup> or 3C<sup>pro</sup> is unlikely.

In poliovirus, 3C<sup>pro</sup> cleaves nearly exclusively between members of a Gln-Gly pair (Hanecak <i>et al.</i>, 1982). However, only nine of the 13 Gln-Gly sites that occur in the poliovirus polyprotein are cleaved. Previous studies have concentrated their efforts on determining the substrate recognition signals that dictate the regulation of processing by 3C<sup>pro</sup>. Signals at the primary amino acid level have been characterized in great detail. It is known that, apart from a specificity for the Gln-Gly dipeptide, flanking amino acids, notably the P<sub>4</sub> position of the scissile bond (the newly generated carboxy terminus of a peptide bond is designated P1, followed by the P<sub>2</sub>, P<sub>3</sub> etc. residues; the newly generated amino terminus is designated P1', followed by the P2', P3' etc. residues) (Hellen <i>et al.</i>, 1989), where an Ala or a small neutral aliphatic amino acid prevails (Nicklin <i>et al.</i>, 1986), contribute to substrate recognition as well as to the efficiency of cleavage (Kuhn <i>et al.</i>, 1988; Blair <i>et al.</i>, 1990). Moreover, there are structural determinants that also carry cleavage recognition signals which may be the dominant determinants in the processing of the capsid precursor. Ypma-Wong & Semler (1987) and Nicklin <i>et al.</i> (1987) have suggested that the entire P1 precursor (Fig. 1) is required for cleavage at Gln-Gly pairs within
Fig. 1. Gene organization, processing scheme and cleavage sites of the poliovirus polyprotein. Proteolytic cleavages of the polyproteins occur between amino acid pairs indicated by standard single-letter code. Arrows above and below the polyprotein indicate sites that are cleaved in cis and trans, respectively, by proteinases as indicated. The question mark indicates that the mechanism of cleavage at this site is not known. The positions of virus-encoded proteinases within the polyprotein are indicated by shaded boxes. Not shown is a ninth Q-G cleavage (within 3C pro) that yields polypeptide P3-4a in small amounts (Nicklin et al., 1986). The nomenclature of poliovirus proteins is according to Rueckert & Wimmer (1984). Reproduced from Hellen et al. (1989).

P1. Furthermore, the proper folding of the P1 β-barrel structure is also required for efficient cleavage at Gln-Gly pairs (Ypma-Wong et al., 1988a).

To extend the study of the substrate specificity of the poliovirus proteinase 3C pro, we engineered 3C pro-directed cleavage sequences into an artificial site, the BC loop of the capsid protein VP1. The BC loop is of particular interest for mutagenesis because (i) it is well exposed on the surface of the virion, because it can bind virus-neutralizing monoclonal antibodies (reviewed in Wimmer et al., 1986), and (ii) it is the only trypsin-sensitive site in all type 2 and type 3 poliovirus strains examined, as well as in the poliovirus type 1 Sabin strain [PV1(S)] (Fricks et al., 1985; Roivanien & Hovi, 1988). Fig. 2 compares the amino acid sequence of the BC loop of PV1(S) and PV1 strain Mahoney [PV1(M)]. The scissile bond cleaved by trypsin in PV1(S) is located between positions 1099 and 1100 (the first digit of these numbers refers to capsid protein VP1, the following numbers to the position of the amino acids in the polypeptide). Interestingly, the two Lys residues, located at positions 1101 and 1103, are not cleaved by trypsin in either PV1(M) or PV1(S) (Fricks et al., 1985).

Using a mutagenesis cartridge containing the cDNA encoding the BC loop of VP1 (Murray et al., 1988), two mutations were introduced in this virion surface element (Fig. 2). The first mutant, W1-1D-BC1, has four amino acids of the wild-type sequence at positions 1099 to 1102 replaced by the amino acid sequence QGPG, thus retaining the native length of the BC loop. This mutation conserves the Ala in the P4 position and places a Pro into the P2' position; a Pro residue frequently occurs in the P2' position of the poliovirus polyprotein (Nicklin et al., 1986) and also occurs in sites that are cleaved efficiently by the 3C pro of encephalomyocarditis virus (Palmenberg et al., 1984). In W1-1D-BC1, the cleavage site (Gln-Gly) for 3C pro is at exactly the same position (amino acids 1099 to 1100) as the Lys-Asn site in PV1(S) that is cleaved by trypsin (Fig. 2). In the second mutant, W1-1D-BC2, a Thr in position 1099 was deleted from the wild-type sequence and a synthetic oligonucleotide encoding nine additional amino acids, GTAKVQGPG, was inserted, of which eight residues (TAKVQGPG) correspond to an authentic 3C pro cleavage site between poliovirus polypeptides 3B and 3C; thus the length of the BC loop was increased to 17 amino acids. Once again, the Ala residue in the P4 position and the Pro residue in the P2' position were conserved.

Fig. 2. Amino acid sequence of PV1(S), PV1(M) and mutants W1-1D-BC1 and W1-1D-BC2 within the mutagenesis cartridge of the BC loop. The mutagenesis cartridge is enclosed in the large box. The BC loop is represented by the amino acids in positions 96 to 104, located between β-strand B and β-strand C (designated by the arrows) (Hogle et al., 1985). Amino acid replacements and insertions in W1-1D-BC1 and W1-1D-BC2 respectively are shown in white as compared to the wild-type PV1(M) and PV1(S) sequences displayed in grey. The scissile bond cleaved by trypsin in PV1(S) is shown by the white diamond (K-N). The scissile bond cleaved by 3C pro in the W1-1D-BC2 virions is shown by the white triangle (Q-G).
Transfection of HeLa cells with RNA transcripts obtained in vitro from linearized full-length cDNA clones of the two mutants resulted in a complete c.p.e. 18 to 36 h post-transfection (van der Werf et al., 1986). From cell lysates, two viable viruses, W1-1D-BC1 and W1-1D-BC2, were recovered, the mutant nucleotide sequences of which were confirmed by sequencing of the viral RNA through the mutagenesis cartridge region. Mutant viruses were further characterized with respect to plaque phenotype and growth kinetics and were shown to be similar, if not identical, to PVI(M) (data not shown). This result is interesting for two reasons. (i) The foreign sequence does not interfere with proper P1 folding, processing and assembly and (ii) the large foreign loop has no apparent effect on the ability of the virus to bind to its receptor.

We were interested in examining whether the newly generated 3C protease cleavage sites in VP1 could be processed by the enzyme. More specifically, we addressed two questions: (i) whether processing of the engineered sites by endogenous 3C protease occurred in vivo; (ii) should in vivo processing not occur, would it be possible to process the cleavage sites in vitro upon incubation with purified 3C protease? For this purpose, 5000 c.p.m. [35S]methionine-labelled, CsCl-purified wild-type and mutant viruses were incubated either with or without purified proteinase A 3C protease isolated after expression in Escherichia coli (Nicklin et al., 1988) for 15 h at 30 °C in the presence of 5 mM-dithiothreitol, 100 mM-NaCl, 20 mM-HEPES, pH 7.5, 1 mM-EDTA. Protein samples were electrophoresed in a 10 to 20% gradient SDS-polyacrylamide gel in the presence of 0.1% SDS (Laemmli, 1970). Fig. 3(a) shows the results of in vivo and in vitro processing by 3C protease at the engineered cleavage sites; no processing of CsCl-purified WI-1D-BC1 virions occurred in vivo (Fig. 3a, lane 4). Similarly, no processing was observed when the virus was incubated in vitro with purified 3C protease at concentrations as high as 25 μM (Fig. 3a, lane 5). A different pattern of processing was seen for the mutant W1-1D-BC2. In the absence of exogenously added 3C protease, CsCl-purified W1-1D-BC2 virions already showed some degree of in vivo processing (although the degree of in vivo processing was variable, it could be as high as 45% when determined by densitometric analyses of autoradiograms), as suggested in lane 6; some of the 33K VP1 protein was processed into 22K and 11K cleavage products. Furthermore, upon incubation of this virus with 25 μM-3C protease, processing of VP1 came close to completion (cleavage was approximately 95% complete, as determined by densitometric analysis of autoradiograms) (Fig. 3a, lane 7). The partial cleavage of VP1 in W1-1D-BC2 viruses isolated from infected HeLa cells may be attributed to in vivo cleavage by endogenous 3C protease or 3CD protease. However, whether this occurred before or after assembly is not known. Moreover, the intracellular concentration of 3C protease or 3CD protease was probably too low to drive cleavage to completion. This is in accord with results indicating the inability of poliovirus-infected
Table 1. Cleavage of synthetic peptides by purified 3Cpro

<table>
<thead>
<tr>
<th>Peptide no.</th>
<th>Sequence</th>
<th>1 h</th>
<th>5 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single substrates*</td>
<td>Ac-MEALFQ*GPLQYKDL-NH2</td>
<td>88.2 ± 0.5</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Ac-PASTQ*GPGKL-NH2</td>
<td>2.0 ± 0.5</td>
<td>8.0 ± 0.7</td>
</tr>
<tr>
<td>Combined substrates§</td>
<td>BIRP0072 Ac-MEALFQ*GPLQYKDL-NH2</td>
<td>91.4 ± 0.2</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>BIRP0313 Ac-PASTQ*GPGKL-NH2</td>
<td>2.2 ± 0.1</td>
<td>9.8 ± 0.3</td>
</tr>
</tbody>
</table>

* Peptides BIRP0072 and BIRP0313 were incubated independently with 12.5 μM-3Cpro at 30 °C.
† Peptide BIRP0072 corresponds to the sequence of amino acids at the in vivo 2C/3A cleavage site of the poliovirus polyprotein.
‡ Peptide BIRP0313 corresponds to the sequence of amino acids in the BC loop of W1-ID-BC1.
§ Peptide BIRP0313 was combined with BIRP0072 in a competition experiment with 12.5 μM-3Cpro at 30 °C. Cleavage was monitored by HPLC 1 h and 5 h post-incubation with enzyme.

HeLa S10 extracts to cleave the BC loop of W1-ID-BC2 virions to completion in vitro (data not shown). W1-ID-BC2 virus was still infectious after what appeared to be nearly complete cleavage of the BC loop by 3Cpro (data not shown). This result is not surprising in the light of data obtained using trypsinized virus (Colbère-Garrapin et al., 1988; Fricks et al., 1985), which is also infectious.

Although terminal amino acid sequencing was not performed to determine the precise cleavage site of this fragment, both cleavage products could be specifically immunoprecipitated with antisera directed against VP1. Fig. 3(b) shows the immunoprecipitation of the cleavage products after incubation with 3Cpro using antisera directed against VP1 and lane 3 shows the specific immunoprecipitation of the 22K and the 11K cleavage products. The specificity of this antiserum is demonstrated in lanes 5 to 7 in which a similar immunoprecipitation experiment was conducted using PVI(M) virions, immunoprecipitating predominantly VP1 (Fig. 3b, lane 6).

The resistance of mutant W1-ID-BC1 to cleavage by 3Cpro could be due to the structural context of the Gln-Gly cleavage site in the BC loop of VP1, or to the sequence of the BC loop, because this did not exactly correspond to any known sequence of a Gln-Gly cleavage site in the poliovirus polyprotein (Nicklin et al., 1986). Therefore, we synthesized a peptide (BIRP0313) of 10 amino acids in length, the sequence of which corresponded exactly to that of the BC loop of VP1 in W1-ID-BC1 (PASTQGPGKL), and tested its susceptibility to cleavage by 3Cpro. The synthesis and purification of peptides were performed as previously described (Pallai et al., 1989). Peptide cleavage assays were performed at 30 °C in the presence of 10 mM-HEPES pH 7.4, 0.1 M-NaCl, 1 mM-EDTA (Pallai et al., 1989). The result of the assay of 3Cpro cleavage of the synthetic peptides is shown in Table 1. Peptide BIRP0313 was cleaved by 3Cpro, albeit at a much lower rate than the test substrate, BIRP0072. As in previous studies, the cleavage products were confirmed to have arisen from scission at the Gln-Gly site (Pallai et al., 1989). Note that cleavage of any suitable peptide substrate with 3Cpro is slow owing to the low turnover rate of the enzyme (M. Skoog & E. Wimmer, unpublished results).

The results presented here strongly suggest that the inability of 3Cpro to cleave mutant W1-ID-BC1, even with a large excess of substrate, is due to structural constraints on the cleavage site. Consequently, by enlarging the BC loop in mutant W1-ID-BC2, we have altered the presentation of the substrate to the binding pocket of the enzyme such that it is now accessible to the active site of the enzyme. This may have been achieved either by relieving structural constraints present in the context of the shorter loop of nine amino acids in wild-type virus, or by increasing the exposure of the cleavage site. The difference of two amino acid residues in the P2 and P3 positions between the engineered cleavage site in W1-1D-BC1 and W1-1D-BC2 virions may also have an effect on cleavage kinetics, although these positions may not exert an effect as strong as that of the P4 position (Nicklin et al., 1986; Pallai et al., 1989).

In this study we have demonstrated also that substrate recognition by proteinase 3Cpro involves sequential determinants as well as structural determinants (Pallai et al., 1989); our results are in agreement with conclusions drawn by Ypma-Wong et al. (1988a). However, these investigators studied the action of endogenous proteolytic activities (i.e. 3Cpro and 3CDpro in infected cell
extracts) on capsid precursors synthesized in vitro. Thus, although processing of folded intermediates occurred, it is possible that processing may not take place in folded end-products. Our approach overcomes this problem by inserting the substrate for cleavage into the intact virion and studying processing by both endogenous enzymatic activity in vivo and purified protease activity in vitro.

The results open up future prospects for the design of substrates specific for protease other than those found in poliovirus by inserting cleavage site-specific sequences in the BC loop of VP1.

We are indebted to Martin J. H. Nicklin for initiating this project, to H. G. Kräusslich for help with construction and characterization of W1-1D-BC2 and to Jonathan Bradely for many valuable suggestions. We are grateful to Peter V. Pallai for providing peptides BIRP0072 and BIRP0313, to Fred Burkhardt for his contribution to the synthesis of peptide BIRP0313 and to Andrew Murdin and Christopher U. T. Hellen for critical reading of the manuscript. Purified 3Cpro was a gift from Kevin Harris and Thomas Hämmerle. This work was supported by the Public Health Service grants AI-15122 and CA-28146 from the National Institutes of Health.

We are grateful to Peter V. Pallai for providing peptides BIRP0072 and BIRP0313, to Fred Burkhardt for his contribution to the synthesis of peptide BIRP0313 and to Andrew Murdin and Christopher U. T. Hellen for critical reading of the manuscript. Purified 3Cpro was a gift from Kevin Harris and Thomas Hämmerle. This work was supported by the Public Health Service grants AI-15122 and CA-28146 from the National Institutes of Health.

References


(Received 15 October 1990; Accepted 25 January 1991)