The protease of adenovirus serotype 2 requires cysteine residues for both activation and catalysis

Alastair W. Grierson, Richard Nicholson, Paul Talbot, Ailsa Webster and Graham Kemp*

Division of Cell and Molecular Biology, School of Biological and Medical Sciences, University of St Andrews, Irvine Building, North Street, St Andrews KY16 9AL, U.K.

Sequence analysis and site-directed mutagenesis were used to study the mechanisms of activation and catalysis of the adenovirus type 2 (Ad2) protease. Primary structure alignments of proteases from 12 serotypes and previously elucidated inhibition profiles were used to target residues for mutagenesis. All conserved serine and cysteine residues were mutated separately and following expression in Escherichia coli their activity in a synthetic peptide assay was compared to that of wild-type recombinant protease. Mutants containing altered serine residues were active while mutations to cysteine-104 and cysteine-122 reduced activity by more than 95%. These results taken together with the known inhibition profile of the adenovirus protease confirm that it is a cysteine protease and suggest that one of these residues provides the active site nucleophile while the other is a part of the thiol–disulphide interchange mechanism previously reported to be involved in its activation.

The production of mature infectious adenovirus type 2 (Ad2) particles requires proteolysis of at least six virus proteins by a virus-encoded protease (Anderson et al., 1973). The mapping of the mutation leading to formation of a temperature-sensitive Ad2 (Ad2ts1) with reduced protease activity at non-permissive temperatures (Weber, 1976; Yeh-Kai et al., 1983), and the demonstration of activity of purified protease using a homogeneous substrate (Webster & Kemp, 1993), have indicated that the viral L3 23K protein is the protease. Protease has been produced from disrupted wild-type virus and has been overexpressed in both E. coli and insect cell (Sf9) systems (Anderson, 1990; Webster et al., 1993). The use of synthetic peptides has shown that the enzyme cleaves at (M,L,I)XGG–X or (M,L,I)XGX–G motifs (Webster et al., 1989b), with the specificity being determined by the four amino acids immediately before the cleavage site. These consensus sequences predict plausible cleavage sites in all of the known substrates and those in the 11K protein and pVI have subsequently been confirmed (Anderson et al., 1989; Anderson, 1990). The sequence of the 23 kDa protein contains no recognizable protease motifs (Anderson, 1990; Houde & Weber, 1990); nor do searches against the databases reveal similarities to any known proteases (Webster et al., 1989a) although the protease sequence is well conserved in the serotypes for which sequence information is available (Cai et al., 1992). Although initially classed as a serine protease (Tremblay et al., 1983), extensive inhibitor studies (Webster et al., 1989a) have suggested that it is a member of the cysteine class, albeit of unusual mechanism, resembling most closely in its inhibition profile the 3C protease of rhinoviruses (Orr et al., 1989).

Recent work has established that the recombinant protease is activated by the peptide GVQSLKRRRCF (pVI-CT), which is identical to the sequence of the C-terminal 11 amino acids of the viral protein pVI (Webster et al., 1993; Mangel et al., 1993). This probably represents the in vitro manifestation of an in vivo mechanism whereby pVI-CT, or more likely pVI itself, is responsible for the regulation of protease activity. It has been shown that the cysteine residue of pVI-CT is essential for the activation process, and it has been postulated that the disulphide bonded dimer of pVI-CT activates the protease by instituting a form of disulphide interchange which either frees the active site thiol or creates an essential stabilizing disulphide bond in the protease or does both (Webster et al., 1993).

In this report we describe the use of sequence analysis and site-directed mutagenesis to confirm that the protease is a cysteine protease and to gain further insight into the catalytic and activation mechanisms.

Protease sequences from 12 adenovirus serotypes were aligned by the initial use of the Pileup program in the University of Wisconsin Genetics Computer Group (GCG) package (Devereux et al., 1984), with an increased gap weight penalty to minimize insertions and deletions in the alignment. Finally the alignment was adjusted manually to bring it into closer accord with the DNA sequence alignment also created with the GCG Pileup
2762 Short communication

Fig. 1. Sections of the alignment of protease sequences from seven human, four animal and one avian adenovirus serotypes, containing the conserved cysteine and serine residues. The alignment was created using the Pileup program in the UWCGG package. Numbering is based on type 2.

program. Relevant sections of this alignment are shown in Fig. 1. This approach produced results in accord with the alignment of proteases from mammalian adenovirus serotypes previously reported by Houde & Weber (1990) in suggesting two conserved serines (positions 95 and 182 using numbering based on the Ad2 sequence). The alignment differed, however, with respect to conserved cysteine residues, indicating two (C104 and C122) rather than the three (C104, C122 and C126) suggested by Houde & Weber (1990). Recently the sequence of the putative protease from an avian adenovirus was reported (Cai & Weber, 1993) and the consequent alignment proposed by these authors also suggests that only the cysteines at 104 and 122 are conserved. Thus if the protease is a cysteine protease then these are the prime candidates for the active site nucleophile.

However, the possibility that the enzyme is a serine protease cannot be ruled out and indeed Houde & Weber (1990) have suggested that serine-160 along with histidine-54 and aspartate-102 constitute the catalytic triad. Site-directed mutagenesis, involving both serine and cysteine has been carried out to clarify the nature of the active site nucleophile. The conserved serine at 182 was changed to an alanine as was cysteine-127 in the Ad2 sequence may substitute. Therefore the mutants C126A, C127A and the double mutant C126A/C127A were all constructed. Following the suggestion that some viral cysteine proteases (in particular those of the picornaviruses) may have an evolutionary relationship with trypsin-like serine proteases, the two fully conserved cysteines were also changed to serines (C104S and C122S).

Site-directed mutagenesis was carried out by two successive PCR reactions using two universal flanking primers and one mutant oligo (Landt et al., 1990). Purified PCR products containing the mutant 23K gene were cloned into the pET11C expression vector (Rosenberg et al., 1987) and transformed into XLI-Blue cells. Plasmid DNA was purified from XLI-Blue cells using QIAGEN Q-20 tips, denatured in 2 M-NaOH, 2 mM-EDTA, and mutations were confirmed by sequencing using the USB Sequenase version 2.0 DNA sequencing kit. The C122A mutant was sequenced in its entirety, while only the regions containing the mutations were sequenced in the others. 23K protein was expressed from transformed BL21(DE3) cells by induction with IPTG. Soluble extracts were prepared as described by Anderson (1990) and applied to a DEFF-Sepharose (Pharmacia) column previously equilibrated with 50 mM-Tris–HCl, pH 8.0. Fractions were collected and protease detected by 15% SDS–PAGE followed by Western blotting with the Amersham ECL kit using antiserum prepared against a synthetic peptide corresponding to the N-terminal 17 amino acids of the Ad2 23K protein (Webster et al., 1993). The concentration of 23K was assessed by comparison of staining intensity with known concentrations of soya bean trypsin inhibitor (Sigma) following SDS–PAGE, using a Vitatron scanning densitometer. Protease activity was assayed in two ways. The first of these was based on the time course of the digestion of the synthetic peptide substrate LSGAGFSW, with cleavage products being analysed and quantified by the Bio-Rad Biofocus 3000 capillary electrophoresis system. Protease (10 to 100 ng) was incubated for 15 min at 37°C with 2 lag pVI-CT in 25 mM-Tris–HCl, 5 mM-EDTA, 1 mM-2-mercaptoethanol before the addition of 20 µg LSGAGFSW. Reaction volume was 100 µl. Aliquots were removed at various time points, the reaction stopped by the addition of an equal volume of 0.1 M-phosphate buffer (pH 2.5). The concentration of GFSW was determined by peak integration following separation in phosphate buffer at 10 kV for 15 min using a 24 cm × 25 µm coated capillary. A standard curve was constructed using synthetic GFSW. The substrate and activating peptides were synthesized by the Fmoc polyamide method of Atherton et al. (1988), purified as described by Webster et al. (1989) and concentrations determined using a fluorescencemine assay (Udenfriend et al., 1972). The second assay involved the use of recombinant preterminal protein (pTP), one of the natural protein substrates of the protease. Purified recombinant 23K or its mutants were incubated with insect cell expressed Ad2 pTP (Webster et al., 1993), monitoring proteolysis by SDS–PAGE followed by Western blotting using both a polyclonal antiserum to pTP (Temperley & Hay, 1991). The results of the protease mutagenesis are summarized in Table 1. Mutation of serines -95, -160 or -182 to an
Table 1. Activities of protease mutants

<table>
<thead>
<tr>
<th>Protease</th>
<th>Relative activity* (%) wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>100</td>
</tr>
<tr>
<td>S95A</td>
<td>110</td>
</tr>
<tr>
<td>S182A</td>
<td>164</td>
</tr>
<tr>
<td>S160A</td>
<td>48</td>
</tr>
<tr>
<td>C104A</td>
<td>3</td>
</tr>
<tr>
<td>C104S</td>
<td>5</td>
</tr>
<tr>
<td>C122A</td>
<td>0.4</td>
</tr>
<tr>
<td>C122S</td>
<td>2</td>
</tr>
<tr>
<td>C126A</td>
<td>62</td>
</tr>
<tr>
<td>C127A</td>
<td>86</td>
</tr>
<tr>
<td>C126A/C127A</td>
<td>103</td>
</tr>
</tbody>
</table>

* The activity of recombinant Ad2 protease was 2100 nmol GFSW/min/mg. Results are from at least three determinations. Standard deviations were ±15% with the exception of S95A and S182A (±20%).

![Fig. 2](image-url)

Fig. 2. Processing of pTP by Ad2 23K protein and its mutants. Purified recombinant pTP (1 µg) was incubated at 37°C for 1 h with 20 µl of FPLC purified fractions containing protease or mutant protease and 2 µg/ml pVI-CT peptide. The assay was done in 25 mM-Tris-HCl, 5 mM EDTA, pH 8.0. Digestion of pTP was detected by 15% SDS-PAGE followed by Western blotting using a polyclonal antiserum to pTP. Cleavage of pTP to iTP is indicated. pTP containing two non-specific degradation products (lane 1) was incubated with: lane 2, wild-type protease+pVI-CT; lane 3, wild-type protease; lane 4, C126A+pVI-CT; lane 5, C104A+pVI-CT; lane 6, C122A+pVI-CT.

alanine residue resulted in a protease which was active. However, when either cysteine-104 or cysteine-122 was changed to alanine or serine, the resultant protease displayed less than 5% of the activity of the wild-type protease. Previous work using thiol-directed inhibitors such as iodoacetate, p-chloromercuribenzoate (pCMB) and N-ethylmaleimide (NEM) (Webster et al., 1989b) had suggested that the Ad2 protease is a cysteine-centre enzyme, and our results using site-directed mutagenesis confirm this to be the case, probably with one of these two residues, which are conserved in all adenovirus serotypes, contributing the active site thiol. The lack of a thiol group at either position 104 or 122 in some way interferes with the protease reaction mechanism either by destroying the active site or by preventing activation via the pVI-CT activating peptide.

When cysteine-126 was changed to alanine (C126A), activity of the resultant protease was not abolished (Table 1). The possibility of cysteine-127 substituting for the mutated cysteine-126 in the catalytic or activation mechanisms was discounted by virtue of the fact that both C127A and C126/127A mutations resulted in active protease preparations. These results suggest that if the activation mechanism involves disulphide exchange exposing an active site thiol as postulated previously (Webster et al., 1993) then one of the other four, non-conserved, cysteine residues must be involved.

Cleavage of recombinant pTP, a natural adenovirus substrate for the protease, is shown in Fig. 2. The behaviour of the protease and its mutants in this assay consolidated the results obtained from the peptide assay. Both wild-type protease and the C126A mutant were clearly shown to digest pTP to the intermediate (iTP) under these assay conditions. The activity of protease with altered residues 104 or 122 was again greatly compromised, with such mutants showing essentially no ability to cleave the pTP substrate. Prolonged exposure of the developed Western blot revealed some iTP in the digests using wild-type protease in the absence of activating peptide (results not shown). This is consistent with the results of Webster et al. (1993) who showed that during coinfection of insect cells with recombinant baculovirus containing the protease gene and recombinant baculovirus containing the DNA coding for the pTP some digestion of the expressed pTP to iTP took place.

The mutants where serines -95, -160 or -182 were substituted with an alanine residue were all active and thus none of these can contribute the active site nucleophile. Given that none of the other serine residues is conserved then the adenovirus protease cannot be a serine protease. Therefore the abolition of activity by mutation of either cysteine-104 or -122, taken along with our previous work (Webster et al., 1989a) showing that it was neither a metallo- nor an aspartyl protease, indicate that the adenovirus enzyme is a cysteine protease.

We have previously reported that the Ad2 protease is activated by the peptide GVQSLKRRRCF and shown that the cysteine residue is an essential component (Webster et al., 1993). The fact that protease activity is abolished by the substitution of either cysteine-104 or cysteine-122 suggests that the activation mechanism also depends on the presence of a critical cysteine residue in the protease. The simplest explanation is that one of the two conserved cysteines provides the active site thiol while the other is essential for interaction with the activating peptide. In an attempt to distinguish which
cysteine is essential for catalysis, both were mutated to serine. In both instances the rate of cleavage of the peptide substrate was greater than with the corresponding alanine mutant, but still less than 5% of the wild-type activity. It would therefore appear that activation and catalysis are inextricably linked. This, plus the observation that the pTP can be digested to iTP by protease in the absence of activating peptide, taken together with the requirement for an approximately 200-fold molar excess of the peptide (Webster et al., 1993) suggests that this peptide does not participate directly in the catalytic mechanism. Therefore activation is more likely to involve the attainment of the appropriate redox state of one or more critical cysteine residues under the influence of the activating peptide.

Recently, Rancourt et al. (1994) have reported results from a similar investigation of the adenovirus protease using site-directed mutagenesis. While both investigations show that the cysteines at positions 104 and 122 are involved in the activation and catalytic mechanisms, the results reported by Rancourt et al., notably the finding that C122G is active, allow the conclusion that cysteine-104 is the active site nucleophile. The sequence surrounding this residue is not that of a conventional protease and although Rancourt et al. (1994) report that there is some residual activity when cysteine-104 is replaced by serine, this activity is abolished by treatment with the thiol-directed reagent iodoacetate, suggesting that the activation and catalytic mechanisms are closely linked. Activation may involve a conformational change which in the wild-type protease is promoted by the interaction of pVI-CT with cysteine-122 and in the C122G mutant is permitted by the small size of the glycine residue.

We are indebted to Jan Armitt for oligonucleotide synthesis and to Bill Blyth for photography. The work was funded by the MRC and R.N. is grateful to the Multiple Sclerosis Society for support. This work benefitted from the use of the SEQNET facility at Daresbury U.K.

References


(Received 1 March 1994; Accepted 30 May 1994)