Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 38 (2004) 108–115
Protein Expression & Purification
www.elsevier.com/locate/yprep

Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro

Sreedevi Nallamsetty[a], Rachel B. Kapust[a], József Tözsér[b], Scott Cherry[a], Joseph E. Tropea[a], Terry D. Copeland[c], David S. Waugh[a,*]

[a] *Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, P.O. Box B, Frederick, MD, United States*
[b] *Department of Biochemistry and Molecular Biology, Research Center for Molecular Medicine, University of Debrecen, Debrecen, Hungary*
[c] *Laboratory of Protein Dynamics and Signaling, Center for Cancer Research, National Cancer Institute at Frederick, P.O. Box B, Frederick, MD, United States*

Received 25 June 2004, and in revised form 13 August 2004

Abstract

Affinity tags are widely used as vehicles for the production of recombinant proteins. Yet, because of concerns about their potential to interfere with the activity or structure of proteins, it is almost always desirable to remove them from the target protein. The proteases that are most often used to cleave fusion proteins are factor Xa, enterokinase, and thrombin, yet the literature is replete with reports of fusion proteins that were cleaved by these proteases at locations other than the designed site. It is becoming increasingly evident that certain viral proteases have more stringent sequence specificity. These proteases adopt a trypsin-like fold but possess an unconventional catalytic triad in which Cys replaces Ser. The tobacco etch virus (TEV) protease is the best-characterized enzyme of this type. TEV protease cleaves the sequence ENLYFQG/S between QG or QS with high specificity. The tobacco vein mottling virus (TVMV) protease is a close relative of TEV protease with a distinct sequence specificity (ETVRFQG/S). We show that, like TEV protease, TVMV protease can be used to cleave fusion proteins with high specificity in vitro and in vivo. We compared the catalytic activity of the two enzymes as a function of temperature and ionic strength, using an MBP–NusG fusion protein as a model substrate. The behavior of TVMV protease was very similar to that of TEV protease. Its catalytic activity was greatest in the absence of NaCl, but diminished only threefold with increasing salt up to 200 mM. We found that the optimum temperatures of the two enzymes are nearly the same and that they differ only two-fold in catalytic efficiency, both at room temperature and 4 °C. Hence, TVMV protease may be a useful alternative to TEV protease when a recombinant protein happens to contain a sequence that is similar to a TEV protease recognition site or for protein expression strategies that involve the use of more than one protease.
Published by Elsevier Inc.

Keywords: Protease; TVMV protease; TEV protease; Tobacco vein mottling virus; Tobacco etch virus; Affinity tag

A major benefit resulting from the advent of recombinant DNA technology has been the large-scale production of proteins of medical or industrial importance. It is common practice to express recombinant proteins in *Escherichia coli* as fusions to highly soluble affinity partners such as maltose-binding protein (MBP) [1] and glutathione *S*-transferase (GST) [2]. This strategy allows for enhanced protein expression levels, greater stability and solubility of the recombinant protein, and provides a convenient affinity handle for protein purification [3,4]. Yet, if these tags remain associated with the protein of interest after purification, they may interfere with protein function, structural analysis, or be immunogenic if the protein is administered in vivo [5]. Removal of the

* Corresponding author. Fax: +1 301 846 7148.
E-mail address: waughd@ncifcrf.gov (D.S. Waugh).

1046-5928/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.pep.2004.08.016

tag is therefore desirable. Frequently used proteases such as thrombin, enteropeptidase, and factor Xa have all been shown to cleave proteins at noncanonical sites [1,6–9]. Consequently, endoproteases with stringent sequence specificity have become important tools [10].

The 3C-like proteases encoded by certain picornaviruses have proven to be the most generally useful reagents. The best-characterized enzymes of this type are rhinovirus 3C protease (also known as PreScission protease) and the nuclear inclusion protease encoded by the tobacco etch virus (TEV), both of which are commercially available [11]. Although these proteases are very sequence specific and rarely, if ever, cleave proteins at noncanonical sites, neither is without drawbacks. PreScission protease has a strict requirement for a Gly–Pro dipeptide in the P1′ and P2′ positions, and so after digestion of a fusion protein with an N-terminal tag a minimum of two nonnative residues will remain on the target protein. TEV protease, on the other hand, has no P2′ specificity (except that Pro is inhibitory in this position) and can tolerate a variety of amino acids in the P1′ position [12]. Yet, TEV protease cleaves itself and autoinactivates [13]. Mutants of TEV protease with greater resistance to autolysis have been described [14,15], but the only mutant that is completely impervious to autoinactivation is also less catalytically active than the wild-type enzyme.

The tobacco vein mottling virus protease (TVMV protease) is closely related to TEV protease (Fig. 1). However, in constrast to the latter enzyme, TVMV protease reportedly does not cleave itself even at high concentrations in solution [16]. Moreover, because its specificity is distinct from that of TEV protease, TVMV protease may be a useful alternative in rare cases when target proteins contain a sequence that happens to resemble the TEV protease recognition site. In the present report, we show that, like TEV protease, TVMV protease can be used to cleave fusion protein substrates with high specificity in vitro and in vivo. The catalytic efficiency and dependence of TVMV and TEV protease activities on monovalent salt concentration and temperature were also compared and found to be similar.

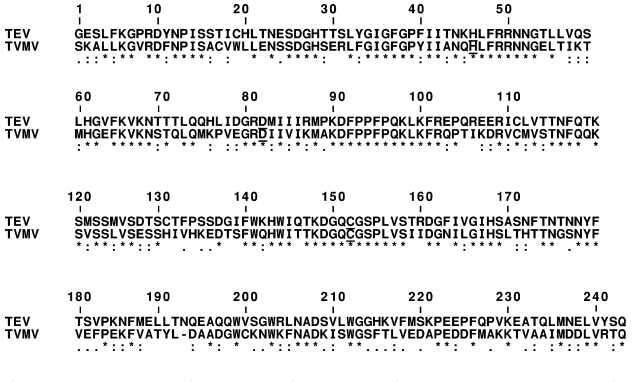

Fig. 1. Sequence alignment of TEV and TVMV proteases. The sequence alignment was made by the program ClustalW. The catalytic triad residues are underlined. Autoproteolysis of TEV protease occurs between Met219 and Ser220.

Material and methods

Construction of plasmid expression vectors

pRK1035

A PCR amplicon carrying an open reading frame (ORF) encoding residues 1–235 of the TVMV protease catalytic domain [16] was constructed from synthetic oligodeoxyribonucleotides, using the program DNA Works [17]. The codon usage was optimized for *E. coli*. In the process, a canonical recognition site for TVMV protease and a His_6-tag (ETVRFQSHHHHHH) were added to the N-terminus of the synthetic TVMV protease ORF. This PCR amplicon was subsequently used as the template for a second PCR with forward and reverse primers that added attB1 and attB2 sites to its N- and C-terminal ends, respectively. The final PCR amplicon was recombined into the destination vector pRK596 [18] to construct the MBP–His_6-TVMV protease fusion vector pRK1035. The complete nucleotide sequence of pRK1035 and additional descriptive information about this plasmid can be obtained at http://mcl1.ncifcrf.gov/waugh_prk1035.html. The MBP–His_6-TVMV protease fusion protein produced by pRK1035 cleaves itself at the designed TVMV recognition site in vivo to yield an N-terminally His_6-tagged TVMV protease that is free of MBP.

pRK1037

The synthetic ORF encoding residues 1–235 of the TVMV protease catalytic domain was amplified from pRK1035 by PCR, using the following oligodeoxyribonucleotide primers: 5′-TCC TCT CAA TTG ATT TAA GGA GGT AAC ATA TGT CTA AAG CTT TGC TGA AGG GC-3′ and 5′-TCT CCT AGA TCT TTA TTA GTC CAT GAT GGC GG-3′. The PCR amplicon was cleaved with *Mfe*I and *Bgl*II, and then ligated with the *Eco*RI/*Bam*HI vector backbone of pRK603 [19] to construct pRK1037, which produces TVMV protease in response to anhydrotetracycline (aTet). The complete nucleotide sequence of pRK1037 and additional descriptive information about this plasmid can be obtained at http://mcl1.ncifcrf.gov/waugh_prk1037.html.

Fusion protein expression vectors

The construction of a Gateway entry clone carrying the *Aquifex aeolicus* NusG–His_6 ORF preceded by a canonical TEV protease recognition site was described previously [12]. An analogous entry clone (pRK1108) in which the NusG–His_6 ORF is preceded by a canonical TVMV protease recognition site (ETVRFQS) instead

was constructed in the same manner. The NusG–His_6 ORF from pRK1108 was recombined into the destination vector pKM596 [18] to construct the MBP–NusG–His_6 fusion protein expression vector pRK1036. The GST–NusG–His_6 fusion vector pRK1109 was constructed by recombining the NusG–His_6 ORF from pRK1108 into the destination vector pDEST3 (Invitrogen). The thioredoxin–NusG–His_6 (TRX–NusG–His_6) fusion vector pSN1606 was constructed in the same fashion, except that the destination vector was pDEST548 (Protein Expression Laboratory, SAIC-Frederick).

Expression and purification of His_6-TEV proteases

Overproduction and purification of wild-type His_6–TEV protease and the autolysis-resistant S219V mutant was performed as described previously [14].

Expression and purification of His_6-TVMV protease

Escherichia coli BL21 CodonPlus-RIL cells (Stratagene) containing pRK1035 were grown to mid-log phase ($A_{600\,nm} = 0.5$) in Luria Broth [20] supplemented with 100 μg/ml ampicillin and 30 μg/ml chloramphenicol at 37 °C, at which time the temperature was shifted to 30 °C and isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to initiate production of the MBP–His_6-TVMV protease fusion protein. After 4 h, the cells were recovered by centrifugation and frozen at −80 °C.

Escherichia coli cell paste was suspended in ice-cold 50 mM sodium phosphate (pH 7.2), 300 mM NaCl, 10% glycerol, and 25 mM imidazole (buffer A) containing Complete EDTA-free protease inhibitor cocktail (Roche Molecular Biochemicals). The cells were lysed with an APV Gaulin Model G1000 homogenizer at 10,000 psi and centrifuged at 30,000*g* for 30 min at 4 °C. The supernatant was filtered through a 0.2 μm polyethersulfone membrane and applied to a 30 ml Ni-NTA Superflow affinity column (Qiagen) equilibrated in buffer A. The column was washed with 10 column volumes of buffer A and eluted with a linear gradient from 25 to 250 mM imidazole in buffer A. Fractions containing recombinant His_6-TVMV protease were pooled and then ethylenediaminetetraacetic acid (EDTA) and dithiothreitol (DTT) were added to final concentrations of 1 and 5 mM, respectively. The sample was concentrated using an Amicon YM10 membrane (Millipore) and diluted 10-fold with 50 mM sodium phosphate (pH 7.2), 10% glycerol, 1 mM EDTA, and 5 mM DTT (buffer B). The sample was applied to a HiPrep 16/10 SP FF column (Amersham Biosciences) equilibrated with buffer B, washed, and eluted with a linear gradient to 250 mM NaCl in buffer B. The peak fractions containing His_6-TVMV protease were pooled and concentrated to 2 mg/ml. Aliquots were flash-frozen with liquid nitrogen and stored at −80 °C until use. The final product was judged to be >95% pure by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). The molecular weight was confirmed by electrospray mass spectrometry.

Expression and purification of MBP–NusG, GST–NusG, and TRX–NusG fusion proteins

Escherichia coli BL21 CodonPlus-RIL cells (Stratagene) containing either pRK1036, pRK1109, or pSN1606 were cultivated and harvested as described above for the production of His_6-TVMV protease. *E. coli* cell paste was suspended in ice-cold 50 mM Hepes (pH 8.0), 200 mM NaCl, and 25 mM imidazole (buffer C) containing complete EDTA-free protease inhibitor cocktail (Roche Molecular Biochemicals). The cells were lysed with an APV Gaulin Model G1000 homogenizer at 10,000 psi and centrifuged at 30,000*g* for 30 min at 4 °C. The supernatant was filtered through a 0.2 μm polyethersulfone membrane and applied to a 10 ml Ni-NTA Superflow affinity column (Qiagen) equilibrated in buffer C. The column was washed and eluted with a linear gradient from 25 to 200 mM imidazole in buffer C. The peak fractions were pooled, concentrated using an Amicon YM10 membrane (Millipore), and then fractionated on a HiPrep 16/60 Sephacryl S-200 HR column (Amersham Biosciences) equilibrated with 50 mM Tris (pH 8), 100 mM NaCl, and 0.5 mM EDTA. Purified recombinant TVMV protease substrates (>90% by SDS–PAGE) were concentrated to 2–4 mg/ml, flash-frozen with liquid nitrogen, and stored at −80 °C until use.

Intracellular processing of fusion proteins

Escherichia coli BL21Pro cells (B & D Biosciences Clontech) containing a fusion protein expression vector (MBP–NusG, GST–NusG, or TRX–NusG) and pRK1037, the TVMV protease expression vector, were grown to saturation in Luria Broth supplemented with 100 μg/ml ampicillin and 30 μg/ml kanamycin at 37 °C. The saturated culture was diluted 1:50 in the same medium and grown to mid-log phase ($A_{600\,nm} = 0.4 - 0.5$) at 37 °C, at which time the temperature was shifted to 30 °C and IPTG was added to a final concentration of 1 mM to initiate production of the fusion protein. To induce the production of TVMV protease, aTet was added to a final concentration of 100 ng/ml 2 h later. After 2 h, the cells from 10 ml of each culture were recovered by centrifugation and resuspended in 1 ml of 50 mM Tris–HCl (pH 8.0), 1 mM EDTA, and 200 mM NaCl. The cell suspensions were lysed by sonication, after which aliquots of the cell lysates were mixed with an equal volume of 2× SDS sample buffer [21] to produce samples of the total protein for SDS–PAGE.

The disrupted cell suspensions were then centrifuged at 14,000*g* for 10 min to pellet the insoluble material. Aliquots of the supernatant fractions were removed and mixed with an equal volume of 2× SDS sample buffer to produce samples of the soluble protein for SDS–PAGE. All samples were heated at 90 °C for 2 min and then centrifuged at 14,000*g* for 5 min prior to SDS–PAGE. Samples were analyzed on 4–12% Bis–Tris NuPage gels (Invitrogen) and visualized by staining with Coomassie brilliant blue.

In vitro processing of fusion proteins

All reactions were performed in 50 mM Tris–HCl (pH 8.0), 0.5 mM EDTA, and 1 mM DTT at 30 °C. The concentration of the fusion protein substrates (MBP–NusG, GST–NusG, and TRX–NusG) was 1 mg/ml. A sample was removed from each reaction for an undigested control, and then TVMV protease was added to the remainder of the reactions at an enzyme:substrate ratio of 1:100 and incubated at 30 °C. Aliquots were removed after 1 h and overnight and mixed with an equal volume of 2× SDS sample buffer. These samples were denatured at 90 °C for 2 min and then centrifuged at 14,000*g* for 5 min prior to SDS–PAGE. The proteins were visualized by staining with Coomassie brilliant blue.

Stability of TVMV vs. TEV proteases against autodigestion

Proteases (either wild-type TEV, S219V TEV, or wild-type TVMV) at a concentration of 1 mg/ml were incubated at room temperature in 50 mM Tris–HCl (pH 8.0), 0.5 mM EDTA and 1 mM DTT. Aliquots were removed at 0 min and after overnight incubation (ca. 12 h). Samples were resolved by SDS–PAGE and proteins were visualized by staining with Coomassie brilliant blue.

Oligopeptide synthesis and characterization

Oligopeptide substrates for TEV and TVMV proteases (TENLYFQSGTRR and TETVRFQSGTRR, respectively) were synthesized by standard 9-fluorenylmethyloxycarbonyl chemistry on a model 430A automated peptide synthesizer (Applied Biosystems) with amide C-terminal ends. Stock solutions were made in distilled water and peptide concentrations were determined by amino acid analysis after peptide hydrolysis, using a Beckman 6300 amino acid analyzer.

Enzyme kinetics

The protease assays were initiated by mixing of 20 µl of TVMV or TEV protease solution in 2× reaction buffer (50 mM sodium phosphate (pH 7.0), 5 mM DTT, 800 mM NaCl, 10% glycerol) with 20 µl of substrate solution (0.04–0.24 mM) in distilled water. The protease solutions were generated by diluting concentrated stock solutions with 2× reaction buffer. The nominal enzyme concentrations used for kinetic measurements were 50 nM TVMV or TEV protease. Higher concentrations (up to 2500 nM) were also tested to confirm that neither enzyme could cleave the canonical recognition site of the other. The actual enzyme concentrations were determined by amino acid analysis. Measurements were performed at six different substrate concentrations. The reaction mixtures were incubated at 30 °C for 30 min and then stopped by the addition of 160 µl 4.5 M guanidine–HCl containing 1% trifluoroacetic acid. An aliquot was injected onto a Nova-Pak C_{18} reversed-phase chromatography column (3.9 × 150 mm, Waters Associates) using an automatic injector. Substrates and cleavage products were separated with an increasing water–acetonitrile gradient (0–100%) in the presence of 0.05% TFA. To determine the correlation between the peak areas of the cleavage products and their amounts, fractions were collected and analyzed by amino acid analysis. The k_{cat} values were calculated by assuming 100% activity for the enzyme. Kinetic parameters were determined by fitting the data obtained at less than 20% substrate hydrolysis to the Michaelis–Menten equation by using the Fig. P program (Fig. P Software). The standard deviations for the k_{cat}/K_M values were calculated according to Boross et al. [22].

Determination of the optimum temperature and salt concentration for TVMV and TEV proteases

Reactions were performed in 50 mM Tris (pH 8.0), 0.5 mM EDTA, and 1 mM DTT at a substrate concentration of 1 mg/ml. Aliquots were removed from each reaction to serve as an undigested control, and then protease was added to the remainder of the reaction to a final concentration of 0.01 mg/ml. Identical reactions were incubated at a series of temperatures (4, 8, 12, 16, 20, 24, 28, 30, 32, 36, and 40 °C). Aliquots were withdrawn at regular intervals (1–10, 20, 30, 40, 50, 60, 120 min, and overnight). Reactions were stopped by the addition of 2× SDS sample buffer, and then the samples were analyzed by SDS–PAGE. Initial velocities were calculated on the basis of data obtained by laser scanning densitometry of Coomassie-stained gels.

The effect of monovalent salt concentration on TEV and TVMV proteases was studied by adding NaCl to standard reaction buffer (above) to achieve a range of salt concentrations (0, 25, 50, 75, 100, 125, 150, 175, and 200 mM NaCl). Reactions were performed at at 30 °C. Aliquots were withdrawn at regular intervals (1–10, 20, 30, 40, 50, 60, 120 min, and overnight), quenched with 2× SDS sample buffer, and subjected to SDS–PAGE. Initial velocities were calculated on the basis of data obtained by laser scanning densitometry of Coomassie-stained gels.

At least three independent experiments were performed to obtain numerical estimates of the fraction of fusion protein that was cleaved at each time point. Coomassie-stained gels were scanned with a Molecular Dynamics Personal Densitometer and the pixel densities of the bands corresponding to the fusion proteins were obtained directly by volumetric integration. The percentage of cleavage at each point was calculated by dividing the amount of fusion protein cleaved by the total amount of fusion protein, after first subtracting the normalized background values obtained from negative control lanes. The mean and standard deviation were generated by Microsoft Excel.

Results

Expression and purification of His-tagged TVMV protease

Expression of TEV protease in *E. coli* was problematic because of rare codons and poor solubility [8,12]. Therefore, to ensure a high yield of TVMV protease in *E. coli*, we constructed a synthetic, codon-optimized gene. To improve the solubility of the enzyme, TVMV protease was produced in the form of a self-processing MBP fusion protein, analogous to the way in which the solubility problem was overcome with TEV protease [8]. A schematic diagram of the TVMV protease expression vector pRK1035 is shown in Fig. 2. The MBP–TVMV fusion protein cleaved itself in vivo to yield an N-terminally His-tagged TVMV protease (ca. 27 kDa) that was partially soluble in the crude cell extract and soluble MBP (ca. 40 kDa) (Fig. 3). The solubility of TVMV protease was somewhat lower than that of TEV protease produced in the same manner [8]. Nevertheless, it was possible to purify ample amounts of the enzyme by immobilized metal affinity chromatography (IMAC) and cation exchange chromatography.

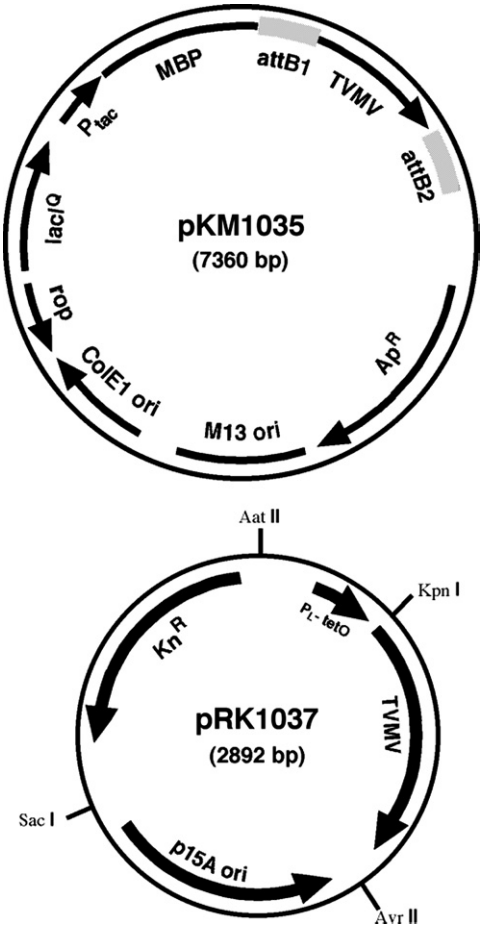

Fig. 2. TVMV protease expression vectors. Schematic diagram (not to scale) of TVMV protease expression vectors pRK1035 and pRK1037.

Digestion of fusion protein substrates in vitro

To gauge how stringent the sequence specificity of TVMV protease is, we digested three fusion protein substrates (MBP–NusG, GST–NusG, and TRX–NusG) in vitro. The three fusion proteins were cleaved efficiently at the designed recognition sites (ETVRFQS), but no cleavage of any of these proteins at noncanonical sites was observed, even after prolonged incubation (Fig. 4). Thus, TVMV protease does not degrade any of these commonly used affinity tags (TRX, GST, MBP) or NusG, suggesting that it is a highly specific enzyme.

Relative stability of TVMV and TEV proteases against autodigestion

Wild-type TEV protease cleaves itself at a specific site near its C-terminus, yielding a truncated enzyme with greatly diminished activity [13]. To confirm that TVMV protease does not digest itself, as previously reported [16], we incubated it under reaction conditions overnight

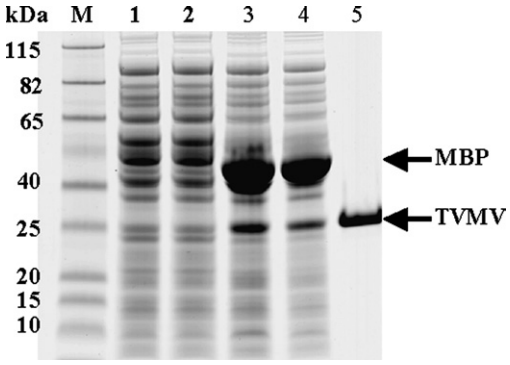

Fig. 3. Expression and purification of MBP–TVMV. Coomassie-stained SDS–PAGE gel (10–20% Tris–glycine gradient, Invitrogen) showing samples of MBP–TVMV under uninduced vs. induced conditions. Lanes: M, broad-range molecular weight standards (kDa); 1 and 2, total and soluble fractions of the uninduced protein, respectively; 3 and 4, total and soluble fractions of the induced protein, respectively; and 5, pure TVMV protease.

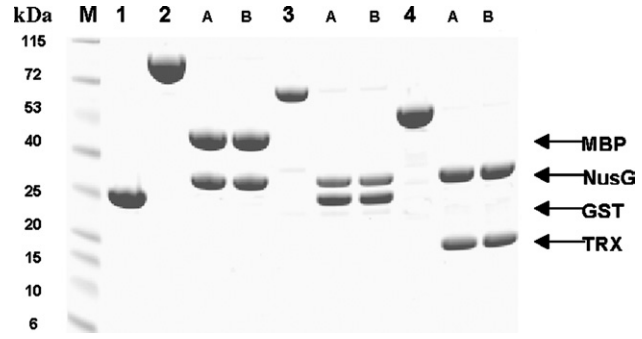

Fig. 4. Digestion of fusion protein substrates in vitro by TVMV protease. Reactions were performed with substrates (MBP–NusG, lane 2; GST–NusG, lane 3; TRX–0NusG, lane 4) at a concentration of 1 mg/ml in 50 mM Tris–HCl (pH 8.0), 0.5 mM EDTA, and 1 mM DTT. Aliquots were removed from each reaction to serve as undigested controls, and then protease was added to the remainder of the reactions to a final concentration of 0.01 mg/ml. The reactions were incubated at 30 °C. Aliquots were removed after 1 h (A) and overnight (B), subjected to SDS–PAGE (10–20% Tris–glycine gradient gel, Novex), and stained with Coomassie brilliant blue. Lane 1 contains pure TVMV protease.

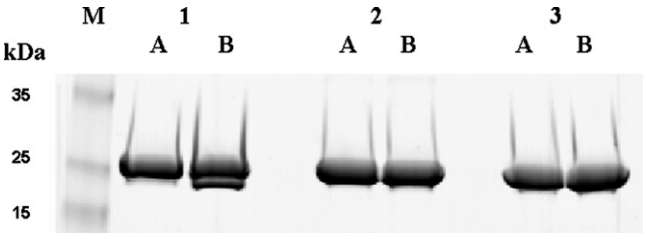

Fig. 5. Stability of TVMV vs. TEV protease against autodigestion. Proteases (wild-type TEV, lane 1; S219V TEV mutant, lane 2; TVMV, lane 3) were incubated at room temperature at a concentration of 1 mg/ml in 50 mM Tris–HCl (pH 8.0), 0.5 mM EDTA, and 1 mM DTT. Aliquots were removed after 0 min (A) and overnight (B). Samples were resolved by SDS–PAGE (10–20% Tris–glycine gradient gel, Invitrogen) and stained with Coomassie brilliant blue.

(Fig. 5) along with wild-type TEV protease and a mutant form of the latter enzyme (S219V) that is resistant to autodigestion [14]. Whereas the wild-type TEV protease cleaved itself under these conditions, both the mutant TEV protease and the wild-type TVMV protease were stable.

Lack of cross-reactivity between TVMV and TEV proteases, and comparison of kinetic parameters

To compare the catalytic activity of TEV and TVMV proteases, kinetic parameters K_M and k_{cat} were determined, using peptide substrates with canonical TEV or TVMV recognition sites (Table 1). The apparent K_M of TVMV protease (0.065 mM) is very similar to that of TEV protease (0.061 mM), but its k_{cat} (0.07 s^{-1}) is half that of TEV protease (0.16 s^{-1}), resulting in an enzyme with ca. twofold lower catalytic efficiency. The K_M and k_{cat} values for TVMV and TEV proteases determined in the present study are similar to those reported previously [12,16,23]. The TVMV and TEV proteases were unable to cleave each other's canonical target sequences, demonstrating that they have distinct sequence specificities.

Comparison of salt and temperature dependence

We wanted to investigate whether TEV and TVMV proteases may have different properties with respect to certain reaction parameters. Two common variables are temperature and salt. As shown in Fig. 6A, the temperature dependence of TVMV and TEV proteases is remarkably similar. Enzymatic activity is maximal at about 30 °C, with a gradual decrease above 34 °C. The decrease in enzymatic activity at higher temperature may

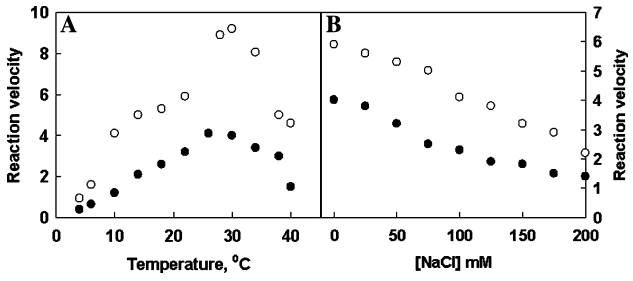

Fig. 6. Determination of the optimal temperature and salt concentration for TVMV and TEV proteases. Reactions were performed in 50 mM Tris (pH 8.0), 0.5 mM EDTA, and 1 mM DTT at a substrate concentration of 1 mg/ml. Aliquots were removed from each reaction to serve as undigested controls and then protease was added to the remainder of the reactions to a final concentration of 0.01 mg/ml. Closed circles and open circles represent TVMV and TEV protease reactions, respectively. (A) Optimal temperature. Identical reactions were incubated at the indicated temperatures and aliquots were withdrawn at regular intervals. The reactions were quenched by the addition of 2× SDS sample buffer and analyzed by SDS–PAGE. Initial velocity was calculated on the basis of data obtained by laser scanning densitometry of Coomassie-stained gels. (B) Optimal salt concentration. Concentrated NaCl was added to the standard reaction buffer (above) to achieve a range of monovalent salt concentrations between 0 and 200 mM as indicated. Aliquots were withdrawn at regular intervals and subjected to SDS–PAGE. Initial velocity was calculated on the basis of data obtained by laser scanning densitometry of Coomassie-stained gels.

Table 1
Kinetic parameters for TVMV and TEV proteases

Substrate	Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($mM^{-1}\,s^{-1}$)
TENLYFQ/SGTRR	TEV	0.061 ± 0.010	0.16 ± 0.01	2.62 ± 0.46
TETVRFQ/SGTRR	TVMV	0.065 ± 0.009	0.07 ± 0.01	1.08 ± 0.17
TENLYFQ/SGTRR	TVMV	Not cleaved		
TETVRFQ/SGTRR	TEV	Not cleaved		

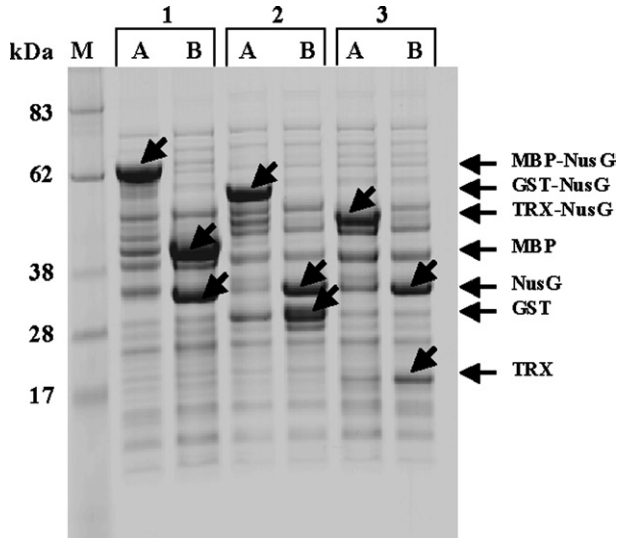

Fig. 7. Intracellular processing of fusion proteins in *E. coli* strain BL21Pro. Cells containing one of the NusG fusion protein expression vectors (MBP–NusG, lane 1; GST–NusG, lane 2; TRX–NusG, lane 3) and pRK1037 were grown to mid-log phase at 37 °C, at which time the temperature was shifted to 30 °C and IPTG was added to initiate the production of the fusion protein (A). To induce TVMV protease expression, aTet was added 2 h later (B). Samples of the total protein were resolved by SDS–PAGE (10–20% Tris–glycine gradient gel, Invitrogen) and stained with Coomassie brilliant blue. Arrows indicate the positions of the fusion proteins and their digestion products.

be due to unfolding of the proteases. Both enzymes are much less active at 4 °C than at their optimal temperatures. Nevertheless, they retain enough activity at 4 °C to cleave a 100-fold molar excess of the MBP–NusG fusion protein overnight (data not shown). The dependence of TVMV and TEV protease activity on NaCl concentration is also similar (Fig. 6B). Maximal activity is obtained at low salt, with a gradual decrease observed as the salt concentration rises.

Intracellular processing of fusion proteins by TVMV protease

In vivo processing of fusion proteins is a convenient way to check processing efficiency and solubility of the released passenger protein before any chromatography experiments are performed [18,19]. pRK1037, which is analogous to the TEV protease expression vector pRK603 [19], can be used to coexpress TVMV protease with fusion protein substrates in *E. coli* (Fig. 2). Three examples of intracellular processing experiments are shown in Fig. 7, using MBP–NusG, GST–NusG, and TRX–NusG fusion proteins as substrates. When cells containing both vectors are induced with IPTG only, little or no processing occurs. This indicates that the TVMV vector is tightly regulated. The ability to regulate the TVMV protease expression vector independently of the IPTG-inducible fusion protein expression vector is an important advantage, because delayed induction of TEV protease often results in greater solubility of the cleaved passenger protein in vivo [18,19]. When cells are induced with both IPTG and aTet, the fusion proteins are efficiently processed in vivo. Although pRK1037 does not produce enough TVMV protease to be visible on the Coomassie-stained gel, enough enzyme is made to cleave virtually all of the overproduced substrates.

Discussion

In addition to their obvious utility for protein purification, affinity tags can improve the yield of recombinant proteins, protect them from intracellular proteolysis, and at least in the case of MBP, enhance their solubility [1,2,7,8]. However, it is ordinarily desirable to remove an affinity tag from a passenger protein for functional and structural studies. Enzymatic methods are most commonly employed to remove affinity tags, yet not all proteases perform this task equally well. While factor Xa, enteropeptidase, and thrombin frequently cleave proteins at noncanonical sites [11,24], TEV protease is highly specific.

Here, we have shown that TVMV protease is similar to TEV protease in terms of its catalytic activity, its sensitivity to monovalent salt concentration, and its temperature-dependence. Both enzymes are active over a wide range of ionic strength and retain appreciable activity at low temperature (4 °C). Like TEV protease, TVMV protease is highly sequence-specific; no nonspecific processing of MBP–NusG, GST–NusG, or TRX–NusG fusion proteins was observed, even at high protease concentrations for extended periods of time. Moreover, unlike TEV protease, TVMV protease does not cleave itself into inactive fragments.

TVMV and TEV proteases have distinct sequence specificities. Neither enzyme was capable of cleaving the canonical recognition site of the other. The most important specificity determinants for TEV protease are EXXYXQS [25] and the canonical target site is considered to be ENLYFQS. An in-depth analysis of TVMV specificity determinants has yet to be performed, but sequence conservation of natural processing sites in the polyprotein suggests that ETVRFQS is the preferred site. V and R are highly conserved in natural TVMV sites and probably are the main discriminators between TVMV and TEV sites. TVMV protease may be a useful alternative to TEV protease when recombinant proteins happen to contain sequence(s) that are similar to the TEV protease recognition site.

It is also possible to imagine situations in which one might wish to take advantage of two proteases with distinct specificities. For instance, a target protein could be produced in the form of a fusion protein with an N-terminal MBP-tag (to take advantage of its solubility-enhancing properties) followed by a biotin acceptor

peptide (BAP). Two distinct protease recognition sites could be placed between the MBP-tag and the BAP and between the BAP and the target protein, respectively. This type of fusion protein could be purified by amylose and/or monomeric avidin affinity chromatography. The MBP-tag could then be removed by one protease, and the resulting BAP-tagged target protein used for surface plasmon resonance experiments or immobilized on avidin/streptavidin-coated surfaces for other purposes without any threat of interference from the MBP-tag. Then, for structural studies, the BAP-tag could be removed by the other protease, yielding a target protein with at most one extra residue appended to its N-terminus. Hence, this strategy would enable multiple objectives to be achieved with a single expression vector.

Acknowledgments

We thank Suzanne Specht for expert technical assistance. Electrospray mass spectrometry experiments were conducted on the LC/ESMS instrument maintained by the Biophysics Resource in the Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute at Frederick.

References

[1] C.V. Maina, P.D. Riggs, A.G. Grandea, E.E. Slatko, L.S. Moran, J.A. Tagliamonte, L.A. McReynolds, C. Guan, A vector to express and purify foreign proteins in *Escherichia coli* by fusion to and separation from, maltose binding protein, Gene 74 (1988) 365–373.

[2] D.B. Smith, K.S. Johnson, Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase, Gene 67 (1988) 31–40.

[3] F. Baneyx, Recombinant protein expression in *Escherichia coli*, Curr. Opin. Biotechnol. 10 (1999) 411–421.

[4] J. Nilsson, S. Stahl, J. Lundeberg, M. Uhlen, P.A. Nygren, Affinity fusion strategies for detection, purification and immobilization of recombinant proteins, Protein Expr. Purif. 11 (1997) 1–16.

[5] D.P. Humphreys, L.M. King, S.M. West, A.P. Chapman, M. Sehdev, M.W. Redden, D.J. Glover, B.J. Smith, P.E. Stephens, Improved efficiency of site-specific copper(II) ion-catalysed protein cleavage effected by mutagenesis of cleavage site, Protein Eng. 13 (1999) 201–216.

[6] T.R. Butt, S. Jonnalagadda, B.P. Monia, E.J. Sternberg, J.A. Marsh, J.M. Stadel, D.J. Ecker, S.T. Crooke, Ubiquitin fusion augments the yield of cloned gene products in *Escherichia coli*, Proc. Natl. Acad. Sci. USA 86 (1989) 2540–2544.

[7] A. Jacquet, V. Daminet, M. Haumont, L. Garcia, S. Chaudoir, A. Bollen, R. Biemans, Expression of a recombinant *Toxoplasma gondii* ROP2 fragment as a fusion protein in bacteria circumvents insolubility and proteolytic degradation, Protein Expr. Purif. 17 (1999) 392–400.

[8] R.B. Kapust, D.S. Waugh, *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused, Protein Sci. 8 (1999) 1668–1674.

[9] L. Rolf, B. Hermann, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements, Nucleic Acids Res. 25 (1997) 1203–1210.

[10] L.M. Babe, C.S. Craik, Viral proteases: evolution of diverse structural motifs to optimize function, Cell 91 (1997) 421–427.

[11] R.C. Stevens, Design of high-throughput methods of protein production for structural biology, Structure Fold Des. 8 (2000) R177–R185.

[12] R.B. Kapust, J. Tozser, T.D. Copeland, D.S. Waugh, The P1′ specificity of tobacco etch virus protease, Biochem. Biophys. Res. Commun. 294 (2002) 949–955.

[13] D.T. Parks, D.E. Howard, T.J. Wolpert, J.D. Arp, G.W. Dougherty, Expression and purification of a recombinant tobacco etch virus NIa proteinase: biochemical analyses of the full-length and a naturally occurring truncated proteinase form, Virology 210 (1995) 194–201.

[14] R.B. Kapust, J. Tozser, J.D. Fox, D.E. Anderson, S. Cherry, T.D. Copeland, D.S. Waugh, Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency, Protein Eng. 12 (2001) 993–1000.

[15] L.J. Lucast, R.T. Batey, J.A. Doudna, Large-scale purification of a stable form of recombinant tobacco etch virus protease, Biotechniques 30 (2001) 544–546.

[16] D.C. Hwang, D.H. Kim, J.S. Lee, B.H. Kang, J. Han, W. Kim, B.D. Song, K.Y. Choi, Characterization of active-site residues of the NIa protease from tobacco vein mottling virus, Mol. Cells 31 (2000) 505–511.

[17] D.M. Hoover, J. Lubkowski, DNA Works: an automated method for designing oligonucleotides for PCR-based gene synthesis, Nucleic Acids Res. 30 (2002) e43.

[18] J.D. Fox, D.S. Waugh, Maltose-binding protein as a solubility enhancer, Methods Mol. Biol. 205 (2003) 99–117.

[19] R.B. Kapust, D.S. Waugh, Controlled intracellular processing of fusion proteins by TEV protease, Protein Expr. Purif. 19 (2000) 312–318.

[20] J.H. Miller, Experimental in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972.

[21] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular cloning: A Laboratory Manual, Second ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

[22] P. Boross, P. Bagossi, T.D. Copeland, S. Oroszlan, L.M. Louis, J. Tözsér, Effect of substrate residues on the P2′ preference of retroviral proteinases, Eur. J. Biochem. 264 (1999) 921–929.

[23] H.Y. Yoon, D.C. Hwang, K.Y. Choi, B.D. Song, Proteolytic processing of oligopeptides containing the target sequences by the recombinant tobacco vein mottling virus NIa proteinase, Mol. Cells 10 (2000) 213–219.

[24] G. Forsberg, B. Baastrup, H. Rondahl, E. Holmgren, G. Pohl, M. Hartmanis, M. Lake, An evaluation of different enzymatic cleavage methods for recombinant fusion proteins, applied on des(1-3)insulin-like growth factor I, J. Protein Chem. 11 (1992) 201–211.

[25] T.D. Parks, K.K. Leuther, E.D. Howard, S.A. Johnston, W.G. Dougherty, Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase, Anal. Biochem. 216 (1994) 413–417.