Protein Engineering vol.14 no.12 pp.993–1000, 2001

Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency

Rachel B.Kapust, József Tözsér[1], Jeffrey D.Fox, D.Eric Anderson[2], Scott Cherry, Terry D.Copeland and David S.Waugh[3]

Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, PO Box B, Frederick, MD 21702-1201, USA and [1]Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Hungary

[2]Present address: Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0580, USA

[3]To whom correspondence should be addressed.
E-mail: waughd@ncifcrf.gov

Because of its stringent sequence specificity, the catalytic domain of the nuclear inclusion protease from tobacco etch virus (TEV) is a useful reagent for cleaving genetically engineered fusion proteins. However, a serious drawback of TEV protease is that it readily cleaves itself at a specific site to generate a truncated enzyme with greatly diminished activity. The rate of autoinactivation is proportional to the concentration of TEV protease, implying a bimolecular reaction mechanism. Yet, a catalytically active protease was unable to convert a catalytically inactive protease into the truncated form. Adding increasing concentrations of the catalytically inactive protease to a fixed amount of the wild-type enzyme accelerated its rate of autoinactivation. Taken together, these results suggest that autoinactivation of TEV protease may be an intramolecular reaction that is facilitated by an allosteric interaction between protease molecules. In an effort to create a more stable protease, we made amino acid substitutions in the P2 and P1′ positions of the internal cleavage site and assessed their impact on the enzyme's stability and catalytic activity. One of the P1′ mutants, S219V, was not only far more stable than the wild-type protease (~100-fold), but also a more efficient catalyst.

Keywords: autoproteolysis/fusion proteins/TEV protease/ tobacco etch virus/

Introduction

Recombinant proteins are often fused to various peptide and protein partners to facilitate their detection and purification, increase their yield and enhance their solubility (Uhlen *et al.*, 1992; Nygren *et al.*, 1994; LaVallie and McCoy, 1995; Nilsson *et al.*, 1997; Baneyx, 1999). Yet, because most affinity tags can be expected to interfere with structural studies and/or the biological activity of the target protein, it is almost always desirable to obtain the native protein free of its fusion partner. Although both chemical and enzymatic methods have been used to cleave fusion proteins at designed sites (Nilsson *et al.*, 1997; Southworth *et al.*, 1999), only the natural proteolytic enzymes have the requisite specificity to be broadly useful reagents for this purpose. The proteases that are most often used are activated blood coagulation factor X (factor Xa), enteropeptidase (enterokinase) and α-thrombin, yet the literature is replete with reports of fusion proteins that were cleaved by these proteases at locations other than the designed site (Forsberg *et al.*, 1991, 1992; He *et al.*, 1993; Wagner *et al.*, 1996; Stevens, 2000).

It is becoming increasingly evident that certain viral proteases have more stringent sequence specificity (Babe and Craik, 1997). These proteases adopt a trypsin-like fold but possess an unconventional catalytic triad in which cysteine replaces serine (Bazan and Fletterick, 1989; Gorbalenya *et al.*, 1989). The nuclear inclusion protease from tobacco etch virus (TEV) is one of the best-characterized enzymes of this type. TEV protease cleaves the amino acid sequence ENLYFQG/S between QG or QS with high specificity (Parks *et al.*, 1994). In contrast to factor Xa, enteropeptidase and thrombin, there have been no published reports of cleavage at non-canonical sites in designed fusion proteins by TEV protease. However, the sequence specificity of TEV protease is not absolutely stringent because all but one of the natural processing sites in the TEV polyprotein deviate from the consensus sequence and TEV protease can cleave many engineered sites that do not exactly match the canonical target site (Dougherty *et al.*, 1988, 1989). Systematic studies have implicated E, Y, Q and, to a lesser extent, G/S, as important specificity determinants (Dougherty *et al.*, 1988, 1989).

Although it has the requisite specificity to be a generally useful reagent for cleaving genetically engineered fusion proteins, a significant shortcoming of TEV protease is that it cleaves itself at a specific site to generate a truncated protease with greatly diminished activity (Parks *et al.*, 1995). Autoinactivation of TEV protease results in substantial losses during purification of the enzyme, and the truncated product is not easy to separate from the full-length protease. Additionally, heterogeneity arising from autoinactivation during the digestion of a fusion protein substrate reduces enzymatic efficiency and complicates the removal of the protease from the digestion products. Moreover, the progressive loss of activity during storage of the enzyme is a considerable nuisance. In an effort to create a more stable but equally active form of TEV protease, we constructed several mutants with single aminoacyl substitutions adjacent to the internal cleavage site and investigated their impact on the stability and catalytic activity of the enzyme.

Materials and methods

His–TEV–Arg proteases

The wild-type and mutant His–TEV–Arg protease catalytic domains used in this study consisted of amino acid residues 189–424 of the mature (49 kDa) nuclear inclusion a (NIa) protease (Dougherty *et al.*, 1989) bracketed by polyhistidine (GHHHHHHH) and polyarginine (RRRRR) sequences on the N- and C-termini, respectively. Residue 1 in our numbering scheme corresponds to residue 189 of the mature 49 kDa TEV NIa protease. Accordingly, the putative catalytic triad residues,

denoted H46, D81 and C151 in our numbering scheme, correspond, respectively, to positions 234, 269 and 339 in the 49 kDa NIa protease. Except for the nucleotides that had to be altered to create the desired amino acid substitutions in TEV protease, all of the His–TEV–Arg constructs used in this study are identical. pRK683, pRK651, pKM607, pRK794, pRK793 and pRK792 encode the wild-type, F217K (TTC to AAA), S219D (AGC to GAC), S219E (AGC to GAA), S219V (AGC to GTG) and S219P (AGC to CCG) His–TEV–Arg proteases, respectively. To construct pRK683, we amplified the open reading frame (ORF) encoding the TEV protease catalytic domain contained on pRK508 (Kapust and Waugh, 1999) by polymerase chain reaction (PCR) with primers PE-29 (5′-GAT GAA GCC CTG AAA GAC GCG CAG-3′) and PE-267 (5′-TAT TAT GGA TCC TTA TTA GCG ACG GCG ACG ACG ATT CAT GAG TTG AGT CGC TTC C-3′). The PCR amplicon was digested with *Sac*I and *Bam*HI, and then ligated with the *Sac*I/*Bam*HI vector backbone fragment of pMal-C2 (New England Biolabs). The mutations were introduced into pRK683 by overlap extension PCR (Ho *et al.*, 1989), using a complementary pair of internal primers containing the desired nucleotide substitutions in concert with PE-29 and PE-30 (5′-GCA AGG CGA TTA AGT TGG GTA ACG C-3′), a pair of external primers that flank the TEV protease ORF. The PCR products were digested with *Sac*I and *Bam*HI, and then ligated between the unique *Sac*I and *Bam*HI sites in pMal-C2 to create the corresponding *Escherichia coli* maltose-binding protein (MBP) fusion vectors. The nucleotide sequence of the insert in each vector was confirmed experimentally. The His–TEV–Arg proteases were generated by intracellular autoprocessing of MBP fusion proteins as described (Kapust and Waugh, 1999), except that the temperature was reduced to 30°C upon addition of isopropyl β-D-thiogalactopyranoside (IPTG). Autoprocessing of the wild-type and mutant fusion proteins proceeded to completion *in vivo* to yield separate MBP and His–TEV–Arg domains.

To purify the wild-type and mutant His–TEV–Arg proteases, a bacterial cell pellet obtained from 6 l of LB medium (Miller, 1972) in shake flasks (typically 18–21 g of wet cell paste) was thawed and resuspended in 200 ml of buffer A: 50 mM sodium phosphate (pH 7.4), 100 mM NaCl. Just prior to lysis, phenylmethylsulfonyl fluoride (PMSF) and benzamidine were added to final concentrations of 1 and 2.5 mM, respectively. After the cells were lysed by sonication, polyethylenimine (Sigma) was added to 0.1% (w/v) and the crude lysate was clarified by centrifugation at 37 000 *g* for 10 min. Solid ammonium sulfate was added to the supernatant to 35% saturation, after which the solution was clarified again by centrifugation as above. The supernatant was then adjusted to 65% saturation with solid ammonium sulfate and incubated on ice for 15 min. The precipitated material was pelleted by centrifugation as above, resuspended in 200 ml of buffer A and filtered (0.45 μM) prior to chromatography. This sample was then applied to a 1.6×20 cm column of Ni-NTA-agarose (Qiagen) pre-equilibrated in buffer A at a flow rate of 2 ml/min. The column was washed with 15 column volumes of buffer A containing 25 mM imidazole, after which the His–TEV–Arg protease was eluted with buffer A containing 200 mM imidazole. The eluted protein was dialyzed for 3 h against 20 volumes of buffer B: 20 mM sodium phosphate (pH 7.4), 50 mM NH_4Cl, 2 mM EDTA, 20 mM 2-mercaptoethanol. After dialysis, the sample was adjusted to 1.1 M ammonium sulfate by slowly adding the solid with constant mixing. The sample was degassed and then applied to a butyl-Sepharose column (1.6×20 cm) pre-equilibrated with buffer B at a flow rate of 3 ml/min. The protein was eluted from the column with three volumes of buffer B containing 1 M ammonium sulfate. The eluted material was dialyzed against 20 mM bicine (pH 8.1), 20 mM NH_4Cl, 1 mM EDTA, 20 mM 2-mercaptoethanol (buffer C) containing 10% (w/v) D-sorbitol until the conductivity had dropped below that of 50 mM NH_4Cl. Because the His–TEV–Arg proteases become highly concentrated during purification, a significant amount of the truncated protease sometimes formed during the first two chromatographic steps, particularly in the case of wild-type His–TEV–Arg. To separate the truncated material from the full-length protease, the dialyzed sample was applied to a column (1.6×10 cm) of Merck Fractogel EMD-COO (20–40 μM) resin, equilibrated with buffer C, at a flow rate of 2 ml/min. The column was washed with 10 column volumes of buffer C. The truncated protease, which did not possess a polyarginine tag, was eluted with buffer C containing 100 mM NH_4Cl. The full-length His–TEV–Arg protease was eluted with buffer C containing 500 mM NH_4Cl. The pure proteases were immediately mixed with an equal volume of glycerol, aliquoted, flash-frozen in liquid nitrogen and stored at –80°C until further use. Using this procedure, we obtained ~2.5 mg of pure His–TEV–Arg protease per gram of wet cell paste.

Catalytically inactive TEV protease

pRK810 directs the expression of a catalytically inactive TEV protease mutant (D81N) bracketed by a biotin acceptor peptide (BAP) on its N-terminus and a His-tag on its C-terminus. This vector was constructed by PCR amplification of the ORF encoding the D81N mutant TEV protease contained on pRK580 (R.Kapust, unpublished data) with primers PE-619 (5′-GTC GGT CTC GAG CGG AGA AAG CTT GTT TAA GGG GCC GCG T-3′) and PE-620 (5′-CTC CCT GGA TCC TTA GTG ATG ATG GTG GTG ATG ATT CAT GAG TTG AGT CGC TTC C-3′). The PCR amplicon was digested with *Xho*I and *Bam*HI and then ligated with the *Xho*I/*Bam*HI vector backbone fragment of pDW363 (Tsao *et al.*, 1996) to create pRK810. The nucleotide sequence of the insert was verified experimentally. BL21/DE3 cells containing pRK810 were grown, induced and harvested as described above for the His–TEV–Arg proteases, except that the medium was supplemented with 50 μM biotin.

To purify the catalytically inactive BAP–TEV(D81N)–His protease, 4 g of cell paste was thawed and resuspended in 40 ml of buffer D: 25 mM HEPES (pH 8.0), 75 mM NaCl, 10% (w/v) D-sorbitol. The cells were lysed with an APV Gaulin G1000 homogenizer at 10 000 psi, after which the insoluble debris was pelleted by centrifugation at 15 000 *g* for 30 min. The supernatant was applied at a flow rate of 2 ml/min to a 1.6×5.0 cm (10 ml) Ni-NTA column that had been equilibrated with buffer D. The column was washed with three volumes of buffer D followed by seven volumes of buffer D containing 30 mM imidazole. The BAP–TEV(D81N)–His protease was subsequently eluted from the column with buffer D containing 100 mM imidazole. Fractions containing the protease were pooled (~50 ml total volume) and further purified by monomeric avidin affinity chromatography, using one-third of the material during each of three successive rounds of chromatography. Each time, ~17 ml of BAP–TEV(D81N)–His protease was applied at a flow rate of 1 ml/min to a 1.6×4.0 cm (7 ml) column containing UltraLink immobilized monomeric avidin (Pierce) that had been equilibrated in buffer

E: 25 mM Tris–HCl (pH 7.5), 25 mM NaCl, 10% (w/v) D-sorbitol. The column was washed with seven volumes of buffer E, after which the biotinylated BAP–TEV(D81N)–His protease was eluted with buffer E containing 2 mM biotin. The column was stripped with four volumes of 0.1 M glycine (pH 2.8) between the chromatography cycles. The fractions containing biotinylated BAP–TEV(D81N)–His protease were pooled and concentrated to 22 ml. This sample was then applied at a flow rate of 3 ml/min to a 2.6×60 cm (320 ml) Sephacryl S-100 HR column equilibrated in buffer F: 20 mM bicine (pH 8.5), 2 mM EDTA, 2 mM 2-mercaptoethanol, 500 mM NH_4Cl. The fractions containing biotinylated BAP–TEV(D81N)–His protease were pooled, concentrated to 2 mg/ml, mixed with an equal volume of glycerol, flash-frozen in liquid nitrogen and stored at –80°C.

Autodigestion experiments

To compare the stability of wild-type and mutant His–TEV–Arg proteases, stock solutions of the enzymes in 50% glycerol were thawed and the proteins were precipitated by the addition of five volumes of saturated ammonium sulfate solution. After 15 min on ice, the precipitates were pelleted by centrifugation at 6000 *g* (10 min at 4°C). The pellets were resuspended in ice-cold TEV protease reaction buffer: 50 mM Tris–HCl (pH 8.0), 0.5 mM EDTA, 1% glycerol, 1 mM dithiothreitol (DTT). The protein concentrations were estimated on the basis of their absorbance at 280 nm (extinction coefficient = 32 410 M^{-1} cm^{-1}), after which the proteases were diluted to the desired concentration (33 μM) with reaction buffer and incubated at 30°C. Aliquots were removed at regular intervals and mixed with three volumes of sample buffer (Laemmli, 1970) to quench the reactions. The 29 kDa His–TEV–Arg proteases and their 26 kDa degradation products were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) on 12% Tris–glycine gels (Novex). Bands were visualized by staining with GelCode Blue (Pierce).

Intermolecular proteolysis was assayed by incubating biotinylated, catalytically inactive BAP–TEV(D81N)–His protease alone or together with an equal concentration (67 μM) of wild-type His–TEV–Arg protease in reaction buffer at 30°C. After 3 h, the reactions were subjected to electrophoresis in a 12% Tris–glycine gel (Novex) and then electrophoretically transferred to a nitrocellulose membrane, using a Novex ExCell II device according to the manufacturer's instructions. The membrane was blocked in Tris-buffered saline [TBS; 50 mM Tris–HCl (pH 7.5), 150 mM NaCl] containing 2% (w/v) bovine serum albumin for 1 h and then washed for another hour in TBS containing 0.2% (v/v) Tween-20 (TBST). Immunopure streptavidin-horseradish peroxidase (Pierce) was added to a final concentration of 0.1 μg/ml and the mixture was incubated overnight at 4°C. The next day, the membrane was washed for 1 h in TBST and developed with SuperSignal chemiluminescence substrate (Pierce).

To detect allosteric effects on autoinactivation, a fixed concentration of the wild-type His–TEV–Arg protease (17 μM) was incubated with increasing concentrations of enzymatically biotinylated, catalytically inactive BAP–TEV(D81N)–His in TEV protease reaction buffer at 30°C. Aliquots were removed from the reactions at various intervals and analyzed by SDS–PAGE as described above.

Oligopeptide synthesis and characterization

Oligopeptides were synthesized by standard 9-fluorenylmethyloxycarbonyl chemistry on an Applied Biosystems model 430A automated peptide synthesizer. The amino acid composition of the peptides was determined with a Beckman 6300 amino acid analyzer. Stock solutions and dilutions were made in distilled water and the peptide concentrations were determined by amino acid analysis.

Enzyme kinetics

The concentration of His–TEV–Arg protease stock solutions was determined by amino acid analysis. The protease assays were initiated by mixing 20 μl of protease solution (40–1200 nM) in 50 mM sodium phosphate (pH 7.0) containing 5 mM DTT, 800 mM NaCl, 10% glycerol, with 20 μl of the peptide SP-3246 (TENLYFQSGTRR-NH_2) in H_2O. The final substrate concentrations ranged between 0.03 and 1.4 mM; the actual range was selected on the basis of the approximate K_m value for each protease. Measurements were performed at six different substrate concentrations in duplicate. The reaction mixtures were incubated at 30°C for 30 min, then stopped by the addition of 160 μl of 4.5 M guanidine hydrochloride (Gd-HCl) containing 1% trifluoroacetic acid (TFA). An aliquot was injected onto a Nova-Pak C18 reversed-phase chromatography column (3.9×150 mm, Waters Associates) using an automatic injector. The substrate and the cleavage products were separated using an increasing water–acetonitrile gradient (0–100%) in the presence of 0.05% TFA. To determine the correlation between peak areas of the cleavage products and their amount, fractions were collected and subjected to amino acid analysis. The k_{cat} values were calculated by assuming 100% activity for the enzymes. 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). Standard deviations for the k_{cat}/K_M values were calculated as described (Boross *et al.*, 1999).

Equilibrium denaturation experiments

Wild-type and mutant TEV proteases were diluted to an A_{280} value of ~0.05, in the presence of 20 mM sodium phosphate (pH 7.5), 5 mM DTT and various concentrations (0–2.8 M) of Gd-HCl. After equilibration at 25°C for 3–4 h, fluorescence spectra were recorded on a FluoroMax-2 spectrofluorometer (Jobin Yvon-Horiba). Samples were excited at 280 nm and emission spectra were recorded between 290 and 500 nm. The average emission wavelengths (Royer, 1995) were plotted as a function of Gd-HCl concentration and fitted to a two-state folding model (Santoro and Bolen, 1988; Pace and Scholtz, 1997) using a non-linear least-squares algorithm. This method of analysis yielded values for the Gd-HCl concentration at the unfolding transition midpoint (C_m) as well as the free energy of unfolding (ΔG°_{H2O}, extrapolated to zero denaturant).

Results

Autoinactivation of TEV protease

Autoproteolysis occurs between Met218 and Ser219 in the TEV protease catalytic domain (Parks *et al.*, 1995), which is peculiar because the surrounding sequence (GHKVFMS) barely resembles the consensus target site (ENLYFQS/G). Only Phe217 and Ser219, which, respectively, occupy the P2 and P1′ positions of the internal cleavage site, match the consensus sequence; suboptimal residues are present in all of the other specificity pockets (Dougherty *et al.*, 1988; Parks *et al.*, 1995). A unimolecular reaction mechanism might explain the relatively facile autolysis of TEV protease at a site that

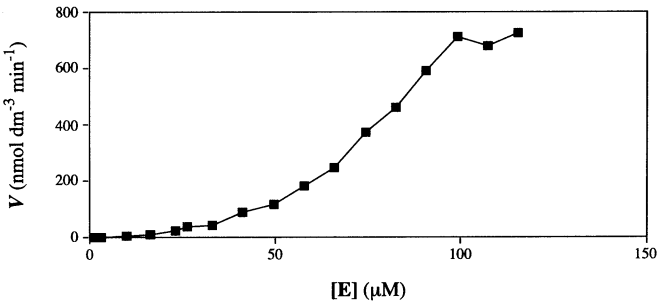

Fig. 1. Concentration dependence of autoinactivation. The wild-type His–TEV–Arg protease catalytic domain was incubated in reaction buffer (see Materials and methods) at various concentrations ranging between 1 and 125 μM (3.5 mg/ml), as indicated. The intact protease and its degradation product were separated by SDS–PAGE and visualized by staining with GelCode Blue. The fraction of protease that was cleaved as a function of time at each concentration was estimated by laser-scanning densitometry of the stained gels, using a Personal Densitometer (Molecular Dynamics). The rate of autoinactivation (V; nmol/dm^3/min) at various protease concentrations [E] was calculated from these data under conditions where less than 20% of the protease was cleaved.

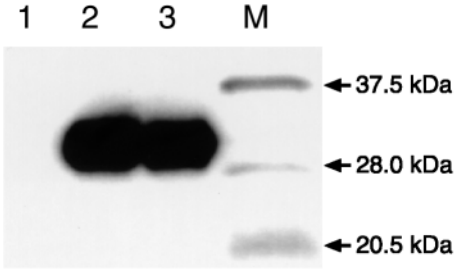

Fig. 2. Direct assay for intermolecular autoinactivation of TEV protease. An enzymatically biotinylated, catalytically inactive TEV protease mutant (D81N) was incubated alone or in combination with an equimolar amount (67 μM) of catalytically active, wild-type His–TEV–Arg protease in reaction buffer for 1 h at 30°C. The reaction products were separated by SDS–PAGE on a 12% Tris–glycine gradient gel, electrophoretically transferred to a nitrocellulose membrane, and then probed with a streptavidin-horseradish peroxidase conjugate. Lanes: 1, His–TEV–Arg only; 2, BAP–TEV(D81N)–His only; 3, His–TEV–Arg + BAP–TEV(D81N)–His. The positions of molecular weight markers (M) are indicated.

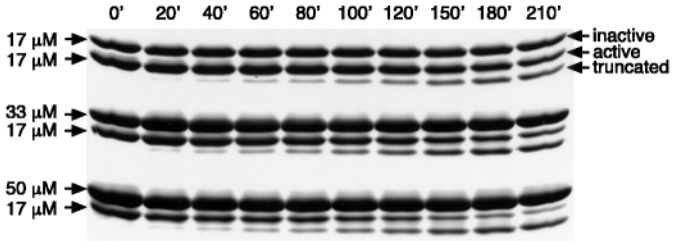

Fig. 3. Allosteric activation of autoproteolysis. A fixed concentration of wild-type His–TEV–Arg protease was incubated in the presence of increasing concentrations of catalytically inactive BAP–TEV(D81N)–His, as indicated, and autodigestion was monitored as a function of time (min) by SDS–PAGE.

bears little similarity to the canonical recognition sequence (Parks *et al.*, 1995). Perhaps the scissile bond is contained within a relatively flexible segment of the protein that is positioned very close to the active site, effectively mimicking a highly concentrated substrate. To investigate this possibility, we examined the concentration dependence of the autoinactivation reaction.

Autoproteolysis of TEV protease was monitored as a function of time, by SDS–PAGE, at a series of protein concentrations ranging between 1 and 120 μM (3.5 mg/ml), and the results were quantified by laser-scanning densitometry of the stained gels. The protease used in this experiment (His–TEV–Arg) consisted of amino acid residues 189–424 of the mature (49 kDa) NIa protease (Dougherty *et al.*, 1989) bracketed by polyhistidine (GHHHHHHH) and polyarginine (RRRRR) sequences on the N- and C-termini, respectively, an arrangement designed to facilitate purification of the enzyme (see Materials and methods). Under pseudo steady-state conditions (i.e. when less than 20% of the protease is cleaved), the data clearly demonstrate a marked concentration dependence of autoinactivation (Figure 1). This behavior is inconsistent with a unimolecular reaction mechanism.

Therefore, we investigated whether autoinactivation of TEV protease occurs via an intermolecular mechanism by incubating a catalytically inactive form of TEV protease, containing a D81N substitution in the catalytic triad (Dougherty and Parks, 1989), with a catalytically active (wild-type) His–TEV–Arg protease. The inactive protease used in this experiment (BAP–TEV(D81N)–His) has a hexahistidine tag on its C-terminus and a BAP on its N-terminus. The latter tag is a substrate for enzymatic biotinylation in *E.coli* (Schatz, 1993). Consequently, if any of the inactive (biotinylated) protease was cleaved by the active His–TEV–Arg enzyme, then we would be able to detect the truncated product with a streptavidin-horseradish peroxidase probe in a western blot experiment. Surprisingly, as shown in Figure 2, the catalytically active His–TEV–Arg protease was unable to convert the catalytically inactive BAP–TEV(D81N)–His protease into the truncated form.

These seemingly paradoxical observations could be reconciled if the mechanism of autoinactivation involves an allosteric interaction between TEV protease molecules that triggers intramolecular proteolysis. To investigate this possibility, we incubated a fixed amount of the wild-type His–TEV–Arg protease (~17 μM) with increasing concentrations of the catalytically inactive BAP–TEV(D81N)–His protein and monitored autoinactivation as a function of time. The active and inactive forms of TEV protease can be distinguished from each other after SDS–PAGE because the BAP tag causes the latter form to migrate more slowly in the gel. As shown in Figure 3, autoproteolysis of wild-type His–TEV–Arg protease was stimulated by the addition of catalytically inactive protease. On the other hand, in accord with our previous observation (Figure 2), no degradation of the catalytically inactive protease was detected in this experiment. An equal concentration of bovine serum albumin failed to stimulate autolysis of wild-type His–TEV–Arg protease, and the addition of catalytically inactive protease neither accelerated nor inhibited the processing of a canonical fusion protein substrate by the wild-type protease (data not shown). We have not been able to detect stable dimers of TEV protease by size-exclusion chromatography under any conditions, and dynamic light scattering measurements conducted under similar experimental conditions (~20–100 μM) yielded a broad, unimodal distibution of apparent molecular weight (data not shown). These observations suggest that interactions between TEV protease molecules are relatively weak and occur only transiently in solution.

F217K and S219D mutants

As stated above, only the residues in the P2 and P1′ positions of the internal cleavage site in TEV protease match the consensus target sequence; the other positions that are known to influence the efficiency of processing (P7, P4 and P1) are

```
      200                                                       242
Ppvna KQWRYNPDEV CWGSLQLKRD IPQSPFTICK LLTDLDG... EFVYTQ
Ppvd  KQWRYNPDEV CWGSLQLKRD IPQSPFTICK LLTDLDG... EFVYTQ
Ppvra KQWRYNPDEV CWGSLQLKRD IPQSPFTICK LLTDLDG... EFVYTQ
Ppvea KQWRYNPDEV CWGSLQLKRD VPQSPFTICK LLTDLDG... EFVYNQ
Pvyn  KSWVYNPDTV LWGPLKLKDS TPKGLFKTTK LVQDLIDH.. DVVVEQ
Pvyhu KSWVYNPDTV LWGPLKLKES TPKGLFKTTK LVQDLIDH.. DVVVEQ
Pemvc KNWKYNPDTV LWGPLQLTKG TPSGMFKTTK MIEDLLAFKS ESVREQ
Tev   SGWRLNADSV LWGGHKVFMS KPEEPFQPVK EATQLMN... ELVYSQ
Tvmv  KNWKFNADKI SWGSFTLVED APEDDFMAKK TVAAIMD... DLVRTQ
Psbmv KGWKYNKNKV CWGGLELLDD NEPEESGLFR MVKLLKSLEE DGVRTQ
Wmv2  KHWFWEPSKI AWGPLNLVEE QPKEEFKISK LVSDLFG... NTVAVQ
                  ENLYFQG
                        S
```

Fig. 4. Alignment of amino acid sequences near the C-termini of potyvirus nuclear inclusion protease catalytic domains. Autodigestion of TEV protease occurs between Met218 and Ser219 (Parks *et al.*, 1995), causing a nearly 10-fold reduction in proteolytic activity (Table I). The canonical TEV recognition site is displayed below the potyvirus sequences and aligned with the internal cleavage site in TEV protease; amino acid residues that are least tolerant of substitution (Dougherty *et al.*, 1989) are underlined. Abbreviations: Ppvna, plum pox virus (isolate NAT); Ppvd, plum pox virus (strain D); Ppvra, plum pox virus (strain Rankovic); Ppvea, plum pox virus (strain El Amar); Pvyn, potato virus Y (strain N); Pvyhu, potato virus Y (strain Hungarian); Pemvc, pepper mottle virus (California isolate); Tev, tobacco etch virus; Tvmv, tobacco vein mottling virus; Psbmv, pea seed-borne mosaic virus; Wmv2, watermelon mosaic virus 2.

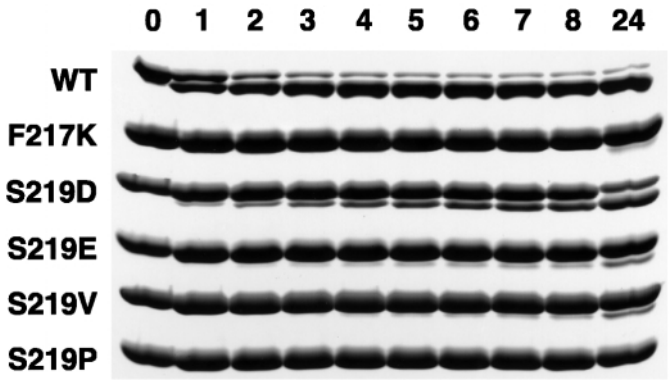

Fig. 5. Autodigestion of wild-type (WT) and mutant TEV protease catalytic domains. The His–TEV–Arg proteases (33 μM, 1 mg/ml) were incubated in reaction buffer for various times ranging between 0 and 24 h, as indicated. The 29 kDa protease and its 26 kDa degradation product were separated by SDS–PAGE on 12% Tris–glycine gradient gels. Bands were visualized by staining with GelCode Blue.

already occupied by unfavorable residues. Accordingly, we reasoned that it might be possible to create a protease that is more resistant to autoinactivation by altering the side chains in the P2 and/or P1′ sites. A rational strategy was suggested by comparing the amino acid sequence of TEV protease with those of some closely related potyviruses (Figure 4). Lys and Asp are the residues that occur most frequently at positions 217 (P2) and 219 (P1′), respectively, and so it seemed reasonable to assume that TEV protease would tolerate these amino acid substitutions. They are also among the least favorable substitutions at these positions in a canonical TEV protease recognition site (Dougherty *et al.*, 1988, 1989). Therefore, we examined the impact of F217K and S219D mutations on the stability and catalytic activity of TEV protease.

The yield and solubility of the F217K and S219D mutants in *E.coli* was similar to that of wild-type TEV protease, but much less truncated material accumulated during the purification of the mutants, particularly the F217K protease (data not shown). The propensity of the mutant proteases to undergo autoinactivation was assessed by incubating them under reaction conditions typically used to digest fusion protein substrates. However, to increase the sensitivity of the experiment, the concentration of protease (1 mg/ml) was 10–20 times greater than would normally be used to cleave a 5 mg/ml solution of a fusion protein. Aliquots were removed at regular intervals over 24 h and the extent of self-processing was monitored by SDS–PAGE. As shown in Figure 5, the majority of the wild-type protease was converted to the truncated form after only 2 h at 30°C. On the other hand, no degradation of the F217K mutant was evident even after 8 h under the same conditions. However, a slight amount of the truncated protease appeared after overnight incubation (24 h) of the F217K mutant. The S219D mutant displayed an intermediate phenotype, undergoing conversion to the truncated form much more slowly than the wild-type protease but far more rapidly than the F217K mutant. Approximately 50% of the wild-type protease was cleaved after 1 h, whereas the S219D mutant required between 8 and 24 h to be degraded to the same extent. Thus, the S219D mutant appears to be approximately 10 times more resistant to autoinactivation than the wild-type protease under these conditions.

Table I. Kinetic parameters for wild-type and mutant TEV proteases[a]

Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
Wild-type	0.061 ± 0.010	0.16 ± 0.01	2.62 ± 0.46
Truncated form	0.448 ± 0.049	0.16 ± 0.01	0.36 ± 0.05
F217K	0.466 ± 0.057	0.15 ± 0.01	0.32 ± 0.05
S219D	0.054 ± 0.011	0.15 ± 0.01	2.78 ± 0.60
S219E	0.074 ± 0.005	0.12 ± 0.01	1.62 ± 0.17
S219V	0.041 ± 0.010	0.19 ± 0.01	4.63 ± 1.16
S219P	0.066 ± 0.008	0.09 ± 0.01	1.36 ± 0.22

[a]With the peptide substrate TENLYFQSGTRR-NH_2.

To ascertain whether either of these mutations interferes with the catalytic activity of TEV protease, we determined the kinetic parameters K_m and k_{cat} for processing of a canonical peptide substrate by wild-type and mutant forms of His–TEV–Arg. The measurements were initially attempted with the peptide substrate SP-4052, having the sequence PTTEN-LYFQSGTVDA-NH_2, which corresponds to that of a naturally occurring processing site in the TEV polyprotein that is cleaved with high efficiency (Dougherty *et al.*, 1988). The peptide was hydrolyzed by the wild-type and mutant enzymes between Gln and Ser, as expected. However, the larger product PTTEN-LYFQ, which had approximately four times the molar extinction coefficient of the smaller product, comigrated with the substrate during HPLC, while the other product partially comigrated with the DTT peak in the chromatogram, making it difficult to determine precise kinetic parameters. To facilitate the separation of the larger cleavage product from the substrate, we synthesized two new peptides: SP-3246 (TENLYFQSG-TRR-NH_2) and SP-3247 (PTTENLYFQSGTRR-NH_2). The addition of two extra Arg residues had a beneficial effect on the solubility of these substrates. SP-3246 and SP-3247 were also cleaved by the proteases between Gln and Ser, with very similar rates, and the large product was separable from the substrate. Kinetic parameters were determined for SP-3246.

The K_m and k_{cat} values for the wild-type and mutant His–TEV–Arg proteases are presented in Table I. All of the enzymes were stable in the assay buffer during the studied time interval,

as monitored by SDS–PAGE (data not shown). The values of K_m and k_{cat} that were obtained for wild-type TEV protease are in excellent agreement with previous results (Parks *et al.*, 1995), indicating that the addition of five arginine residues to the C-terminus of the catalytic domain does not interfere with its activity. The activity of the S219D mutant was indistinguishable from that of the wild-type protease. In contrast, the kinetic parameters for the F217K mutant differed substantially from those of wild-type TEV protease and the S219D mutant, and were remarkably similar to the values obtained for the truncated form of the protease. Our results indicate that both the F217K mutation and the truncation dramatically reduce the affinity of the enzyme for substrate but do not affect the catalytic rate constant. Thus, these kinetic experiments revealed that the F217K mutation has a deleterious effect on the catalytic activity of TEV protease, due to a K_m effect, which may account for its greater degree of resistance to autoinactivation than the S219D mutant.

Additional P1′ mutants

Although it is highly resistant to autoinactivation, the F217K mutant has a severe catalytic defect. Conversely, the S219D mutant exhibits wild-type catalytic efficiency but still undergoes autoinactivation at an appreciable rate. Thus, neither mutation offers a compelling advantage relative to the wild-type protease. Because the F217K substitution drastically impaired the catalytic efficiency of the protease, we were reluctant to make any more mutations at the P2 position. The P1′ site, on the other hand, seemed to be more tolerant of substitution. A systematic analysis of the P1′ specificity of TEV protease revealed that Glu (which is very similar to Asp), the β-branched hydrophobic residues (Leu, Ile and Val), and especially Pro are the most unfavorable residues in this position (R.Kapust *et al*., unpublished observations). We reasoned that if the amino acid sequence in the immediate vicinity of the internal cleavage site contributes to enzyme–substrate recognition in a manner that is consistent with the known specificity of TEV protease, then replacing Ser219 with Glu, Val or Pro should give rise to proteases with greater resistance to autoinactivation than the S219D mutant. Therefore, we constructed and characterized His–TEV–Arg proteases containing these three mutations (S219E, S219V and S219P).

As expected, all three of these mutants proved to be far more resistant to autoinactivation than the S219D protease (Figure 5). The S219E and S219V proteases were converted to the truncated form at a rate that is ~10-fold slower than the S219D mutant (compare S219E and S219V after 24 h with S219D after 2 h) or ~100-fold slower than the wild-type protease. In fact, the S219E and S219V mutants were only slightly less resistant to autolysis than the F217K mutant, and the S219P mutant appeared to be virtually impervious to autoinactivation under all conditions. The values of K_m for the S219E, S219V and S219P mutants were essentially the same as that of the wild-type protease (Table I). The k_{cat} of the S219V mutant was somewhat better than that of the wild-type protease, whereas the corresponding values for the S219E and S219P mutants were ~50–80% that of wild-type His–TEV–Arg. A representative plot of initial velocity versus substrate concentration for one of the mutant proteases (S219P) is shown in Figure 6.

Global stability of mutant TEV proteases

Because the mutant proteases were intended to be used as reagents for cleaving genetically engineered fusion proteins *in vitro*, we thought it would be prudent to compare the global stability of the S219 mutants with that of the wild-type protease in an equilibrium denaturation experiment monitored by fluorescence spectrometry (Figure 7). All of the proteases exhibited a sigmoidal, two-state unfolding transition, with a midpoint Gd-HCl concentration near 1.2 M and ΔG°_{H2O} on the order of 5–7 kcal/mol. Thus, none of the mutations at the S219 position adversely affect the stability of TEV protease.

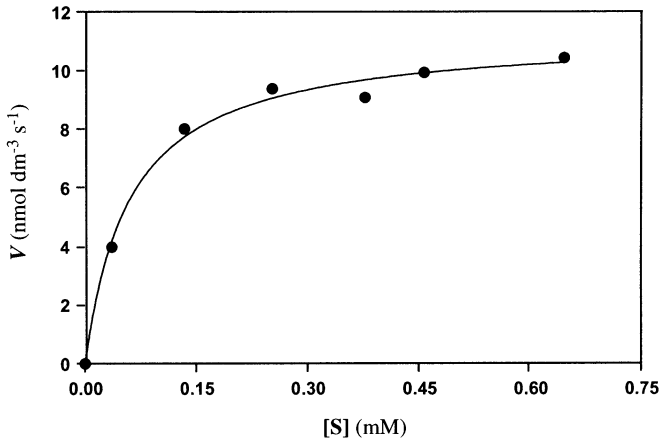

Fig. 6. Kinetic analysis of the TEV protease mutant S219P. Initial velocities at six substrate concentrations were fitted to the Michaelis–Menten equation. The concentration of protease in the reactions was 125 nM.

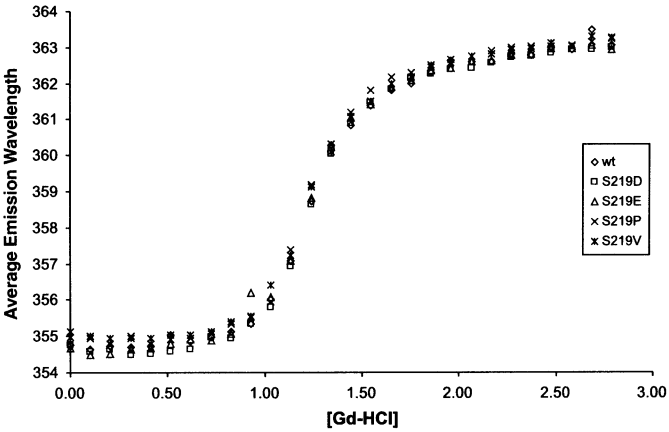

Fig. 7. Equilibrium denaturation of wild-type (wt) and mutant TEV proteases. The weighted average emission wavelengths (nm) for each sample are plotted as a function of Gd-HCl concentration (M).

Discussion

In an effort to reduce or abolish autoproteolytic inactivation of TEV protease, we constructed and characterized several mutants with amino acid substitutions adjacent to the internal cleavage site. This general strategy had been used successfully in the past to inhibit the autoinactivation of other proteases (Rose *et al.*, 1993; Dang *et al.*, 1996; Laco *et al.*, 1997; Tomasselli *et al.*, 1998; Louis *et al.*, 1999; Pray *et al.*, 1999). The amino acid residues corresponding to the P2 and P1′ sites were targeted for mutagenesis because these were the only positions that matched the canonical (consensus) TEV protease recognition site; suboptimal residues were present in the other positions. All of the mutant TEV proteases exhibited a greatly reduced propensity for autoinactivation. Thus, our results demonstrate that the amino acid sequence in the immediate vicinity of the internal cleavage site contributes to enzyme–

substrate recognition in a manner that is consistent with the known specificity of TEV protease, and lend further support to the notion that knowledge about the specificity of a protease can be used to guide the engineering of mutants with a reduced propensity for autoproteolysis (Rose *et al.*, 1993).

An amino acid substitution in the P2 position of the internal cleavage site (F217K) virtually eliminated autoinactivation under normal reaction conditions. Unfortunately, the catalytic activity of the F217K mutant was seriously impaired, effectively nullifying its utility as a reagent. The impact of the F217K mutation was manifest on K_m but not on k_{cat}, which suggests that this side chain contributes, either directly or indirectly, to substrate binding. In contrast, amino acid substitutions at the nearby P1′ position had little or no effect on the catalytic activity of the protease, nor did they significantly reduce the global stability of the protein. However, the S219D mutation was only moderately effective at inhibiting autoinactivation of TEV protease. On the other hand, the S219E and S219V mutants were both highly resistant to autoinactivation (Figure 5) and the S219P mutant appeared to be virtually impervious to autolysis under all conditions tested. The latter mutant exhibited only a 2-fold reduction in k_{cat} and the kinetic parameters determined for the S219V mutant were as good or better than those of the wild-type protease. Thus, the S219V and S219P mutants seem ideally suited for use as proteolytic reagents to cleave genetically engineered fusion proteins.

During the course of this work, we made some interesting observations pertaining to the mechanism of autoinactivation. In contrast to a previous study (Parks *et al.*, 1995), we found that the rate of autoinactivation is proportional to the concentration of TEV protease. This discrepancy is probably due to the fact that the highest concentration examined by Parks *et al*. (~8 μM) corresponds to the lowest part of the curve in Figure 1, where the concentration dependence is far less pronounced. The simplest explanation for the concentration effect we observed is that autoinactivation is the result of intermolecular proteolysis. Therefore, we were surprised to discover that the wild-type enzyme was unable to cleave a catalytically inactive form of the protease (Figure 2). We cannot rule out the possibility that the conservative amino acid substitution in the active site (D81N) distorts the structure of the inactive protease enough to render it refractory to autoprocessing, but this would not explain why the catalytically inactive protease stimulated autoinactivation of the wild-type enzyme in a concentration-dependent manner. An alternative interpretation of our results is that autolysis of TEV protease is an intramolecular event that is stimulated by an allosteric interaction between protease molecules, reminiscent of the manner in which *E.coli* RecA stimulates autolysis of the LexA repressor (Little, 1984). Further research will be necessary to confirm this hypothesis.

It has been suggested that autoinactivation of TEV protease may play a role in the physiology of viral infection (Parks *et al.*, 1995; Kim *et al.*, 1996). In this regard, it is intriguing to note that autoinactivation also occurs in some related proteases, including turnip mosaic virus protease (Kim *et al.*, 1996) and hepatitis A virus 3C protease (Gauss-Muller *et al.*, 1991), albeit at different locations. On the other hand, the closely related tobacco vein mottling virus protease evidently does not undergo autoinactivation *in vitro* (Hwang *et al.*, 2000), so this does not appear to be a universal property of potyviral proteases. If, as we have suggested, autoinactivation is triggered by an allosteric interaction between TEV protease molecules, then a regulatory role seems plausible. Very little is known about the regulation of NIa protease during viral infection, but we note that both the 27 kDa catalytic domain and its 25 kDa autolysis product, which correspond, respectively, to the full-length and truncated His–TEV–Arg proteases analyzed in this study, have been detected in preparatons of NIa protein isolated from infected tobacco plants and have also been observed to accumulate when the 49 kDa NIa precursor is synthesized in a cell free transcription/translation system (Dougherty and Parks, 1991; Parks *et al.*, 1992, 1995). The autoinactivation-resistant mutants described here could be used to clarify what role, if any, autoinactivation plays in the regulation of TEV protease activity during viral replication.

Acknowledgements

We thank Karen Routzahn for constructing pKM607 and pKM773, Suzanne Specht for assistance with peptide synthesis and amino acid analysis and Anne Arthur for expert editorial advice.

References

Babe,L.M. and Craik,C.S. (1997) *Cell*, **91**, 427–430.
Baneyx,F. (1999) *Curr. Opin. Biotechnol.*, **10**, 411–421.
Bazan,J.F. and Fletterick,R.J. (1989) *Virology*, **171**, 637–639.
Boross,P., Bagossi,P., Copeland,T.D., Oroszlan,S., Louis,J.M. and Tozsér,J. (1999) *Eur. J. Biochem.*, **264**, 921–929.
Dang,L.C., Talanian,R.V., Banach,D., Hackett,M.C., Gilmore,J.L., Hays,S.J., Mankovich,J.A. and Brady,K.D. (1996) *Biochemistry*, **35**, 14910–14916.
Dougherty,W.G. and Parks,T.D. (1989) *Virology*, **172**, 145–155.
Dougherty,W.G. and Parks,T.D. (1991) *Virology*, **183**, 449–456.
Dougherty,W.G., Carrington,J.C., Cary,S.M. and Parks,T.D. (1988) *EMBO J.*, **7**, 1281–1287.
Dougherty,W.G., Cary,S.M. and Parks,T.D. (1989) *Virology*, **171**, 356–364.
Forsberg,G., Brobjer,M., Holmgren,E., Bergdahl,K., Persson,P., Gautvik,K.M. and Hartmanis,M. (1991) *J. Protein Chem.*, **10**, 517–526.
Forsberg,G., Baastrup,B., Rondahl,H., Holmgren,E., Pohl,G., Hartmanis,M. and Lake,M. (1992) *J. Protein Chem.*, **11**, 201–211.
Gauss-Muller,V., Jurgensen,D. and Deutzmann,R. (1991) *Virology*, **182**, 861–864.
Gorbalenya,A.E., Donchenko,A.P., Blinov,V.M. and Koonin,E.V. (1989) *FEBS Lett.*, **243**, 103–114.
He,M., Jin,L. and Austen,B. (1993) *J. Protein Chem.*, **12**, 1–5.
Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) *Gene*, **77**, 51–59.
Hwang,D.C., Kim,D.H., Kang,B.H., Song,B.D. and Choi,K.Y. (2000) *Mol. Cells*, **10**, 148–155.
Kapust,R.B. and Waugh,D.S. (1999) *Protein Sci.*, **8**, 1668–1674.
Kim,D.H., Hwang,D.C., Kang,B.H., Lew,J., Han,J., Song,B.D. and Choi,K.Y. (1996) *Virology*, **226**, 183–190.
Laco,G.S., Fitzgerald,M.C., Morris,G.M., Olson,A.J., Kent,S.B. and Elder,J.H. (1997) *J. Virol.*, **71**, 5505–5511.
Laemmli,U.K. (1970) *Nature (London)*, **227**, 680–685.
LaVallie,E.R. and McCoy,J.M. (1995) *Curr. Opin. Biotechnol.*, **6**, 501–506.
Little,J.W. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 1375–1379.
Louis,J.M., Oroszlan,S. and Tozser,J. (1999) *J. Biol. Chem.*, **274**, 6660–6666.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Nilsson,J., Stahl,S., Lundeberg,J., Uhlen,M. and Nygren,P.A. (1997) *Protein Expr. Purif.*, **11**, 1–16.
Nygren,P.A., Stahl,S. and Uhlen,M. (1994) *Trends Biotechnol.*, **12**, 184–188.
Pace,C.N. and Scholtz,J.M. (1997) *Protein Structure: A Practical Approach*. IRL Press, Oxford, UK, pp. 299–321.
Parks,T.D., Smith,H.A. and Dougherty, W.G. (1992) *J. Gen. Virol.*, **73**, 149–155.
Parks,T.D., Leuther,K.K., Howard,E.D., Johnston,S.A. and Dougherty,W.G. (1994) *Anal. Biochem.*, **216**, 413–417.
Parks,T.D., Howard,E.D., Wolpert,T.J., Arp,D.J. and Dougherty,W.G. (1995) *Virology*, **210**, 194–201.
Pray,T.R., Nomura,A.M., Pennington,M.W. and Craik,C.S. (1999) *J. Mol. Biol.*, **289**, 197–203.
Rose,J.R., Salto,R. and Craik,C.S. (1993) *J. Biol. Chem.*, **268**, 11939–11945.
Santoro,M.M. and Bolen,D.W. (1988) *Biochemistry*, **27**, 8063–8068.

Schatz,P.J. (1993) *Biotechnology (NY)*, **11**, 1138–1143.
Southworth,M.W., Amaya,K., Evans,T.C., Xu,M.Q. and Perler,F.B. (1999) *Biotechniques*, **27**, 110–120.
Stevens,R.C. (2000) *Struct. Fold. Des.*, **8**, R177–R185.
Tomasselli,A.G., Paddock,D.J., Curry,K.A., Garlick,R.L., Leone,J.W., Lull,J.M. *et al.* (1998) *Protein Expr. Purif.*, **14**, 343–352.
Tsao,K.L., DeBarbieri,B., Michel,H. and Waugh,D.S. (1996) *Gene*, **169**, 59–64.
Uhlen,M., Forsberg,G., Moks,T., Hartmanis,M. and Nilsson,B. (1992) *Curr. Opin. Biotechnol.*, **3**, 363–369.
Wagner,C.R., Bergstrom,C.P., Koning,K.R. and Hanna,P.E. (1996) *Drug Metab. Dispos.*, **24**, 245–253.

Received May 13, 2001; revised August 25, 2001; accepted September 13, 2001